

STIC-Biotech/ChemLib

95965

RECEIVED

JUN -6 2003

(STIC)

From: Ibrahim, Medina A.
Sent: Thursday, June 05, 2003 5:39 PM
T: STIC-Biotech/ChemLib
Subject: 10/090, 035

Please search the following:

1. DNA encoding SEQ ID NO:2 or 4.
2. SEQ ID NO:1-4.
3. oligo of SEQ ID NO:1 and 3

Please search both commercial and patent databases, including pending. Thanks

Medina A. Ibrahim
Patent Examiner
GAU:1638
CM1-9E03
mailbox-9E12
(703)306-5822

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/6/03
Date Completed: 6/15/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: 4
NA Sequences: _____
AA Sequences: 2+2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

reverse to NA
VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 01/05
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library, CM1 - Circ-Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 08:47:15 ; Search time 1860 Seconds
(without alignments)
1455.140 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYQEVQDYCVSEVRVAPA.....INTCGVHERRESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q/cgn2_1/USPTO_spool/US10090035/runat_06062003_105503_10957/app_query.fasta_1.526
-DB-GenEmbl -QFWT-fastap -SUFEL-rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT-pto -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US10090033@cgn_1_1_2566-brunat_06062003_105503_10957 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*

- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	295	57.8	173699	8	AP003416 Oryza sat
C 2	295	57.8	174218	2	AC091071 Oryza sat
C 3	295	57.8	185095	8	AP003436 Oryza sat
C 4	295	57.8	185481	2	AC091088 Oryza sat
C 5	225.5	44.2	562	8	SSW242802
C 6	93.5	18.3	300000	1	SMES91784
C 7	92.5	18.1	814	8	AF003103
C 8	92.5	18.1	815	8	ATAJ2598
C 9	92.5	18.1	1009	8	AY057683
C 10	92.5	18.1	10757	1	AE004561
C 11	92.5	18.1	198780	8	ATCHRIV86
C 12	92.5	18.1	206420	8	ATAF21
C 13	92	18.0	1898	10	MUSDELTA
C 14	92	18.0	2330	10	MUSUCRBP
C 15	92	18.0	3041	10	MUSUCRBP
C 16	91.5	17.9	149879	2	AP005381
C 17	91.5	17.9	163285	2	AP004585
C 18	91	17.8	130814	2	AC121354
C 19	91	17.8	177392	2	AC127187
C 20	90.5	17.7	203050	1	AL646071
C 21	89	17.5	82972	8	NCH24H17
C 22	88.5	17.4	992	10	BC027007
C 23	88.5	17.4	158414	10	AC005960
C 24	88.5	17.4	196050	1	AL646058
C 25	88.5	17.4	197355	2	AC116130
C 26	88	17.3	8874	9	HSREP10
C 27	88	17.3	197050	1	AL646081
C 28	87.5	17.2	163660	2	AC100745
C 29	87	17.1	110000	2	LMFLCHR16_02
C 30	87	17.1	167050	2	AC119146
C 31	87	17.1	171749	2	AC112440
C 32	87	17.1	178665	2	AP003483
C 33	87	17.1	179137	2	AP000853
C 34	87	17.1	184569	9	AC010583
C 35	87	17.1	210872	10	AC084386
C 36	86.5	17.0	1458	10	MUSBDL2A
C 37	86.5	17.0	1766	10	BC016434
C 38	86.5	17.0	4151	10	AF180338
C 39	86.5	17.0	37805	3	AF125964
C 40	86.5	17.0	181419	10	AL591665
C 41	86	16.9	4131	9	AB023233
C 42	86	16.9	63391	8	NC68B2
C 43	86	16.9	147481	2	AC067762
C 44	86	16.9	165120	9	AL138704
C 45	86	16.9	192149	9	AC091884

ALIGNMENTS

AP003416/c	173699 bp	DNA	linear	PLN 17-MAY-2002
LOCUS				
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0456E05.			
ACCESSION	AP003416			
VERSION	AP003416.4	GI:20804922		
KEYWORDS	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0456E05.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
REFERENCE	1			
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.			
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0456E05			
JOURNAL	Published Only in Database (2001)			
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-MAR-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan			
	(E-mail:tsasak@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)			
COMMENT	On May 15, 2002 this sequence version replaced gi:18181981. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI Nonredundant protein databases with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0456E05 clone has an overlap with P0470A12 (DBJ: AP003436) at the position 130,058 to 173,699 of 3' end. Detailed information on overlap and assembly quality together with an annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.			
FEATURES	Location/Qualifiers			
source	1..173699			
	/organism="Oryza sativa (japonica cultivar-group)"			
	/cultivar="Nipponbare"			
	/db_xref="taxon:39947"			
	/chromosome="1"			
gene	/clone="P0456E05"			
	Join(9483..9641,10723..10785)			
	/gene="P0456E05.1"			
CDS	Join(9483..9641,10723..10785)			
	/gene="P0456E05.1"			
	/note="hypothetical protein"			
	/codon_start=1			
	/protein_id="BAB92602.1"			
	/db_xref="GI:20804923"			
	/translation="MCLMPLSTACVCRQQTQNDWLATPPAPAPAAAALYYIYVLASCDONOCVHARWRVWASIIIN"			
gene	complement(join(11701..11960,11998..12076,12765..12881,13131..13310,13459..13587,13660..13758,14527..14631,15002..15131,15421..15487,17632..17717,17924..17952))			
	/gene="P0456E05.2"			
	complement(join(11701..11960,11998..12076,12765..12881,13131..13310,13459..13587,13660..13758,14527..14631,15002..15131,15421..15487,17632..17717,17924..17952))			
CDS				

/gene="P0456E05.2"	
/note="contains ESTs AU0101913(S0657),AU0101912(S0657) similar to rabgap domains protein"	
unknown protein"	
/codon_start=1	
/protein_id="BAB92603.1"	
/db_xref="GI:20804924"	
/translation="MHEPLYRVHKKOOPVTAQHRPLRGLEATLGSIRRRRSRHETGGFNETVWALLKHEFDCSLWVGEALQSLKVKVLDLDELMLAAGVDPGAARPTVKKLLGLVPSDRALWEQELAKRKSQYAAKEEFLSNPYSEIIEQIDRDVKRTHPDMHFPCDSSEAKSNQESLKNILIEFAKLNAGIRYVQGMNEILAPLFFVFRNDPDKNANFAEADFFCFMELLSGFRDNFCQKLDNSAVGIQGTLSKLSOLVAKYDCELORYLEITFEI NPOFIAPWTITLLITQEFNADTIHNDILLSDPDQPGRSIYRDGTGSMYNTIETS VAHAYQSSRNSAVPVVEKLHRCRKATGAARQAAPFLSPFRSPETTENAFAFPC RFLAATRPISVQGDNLNLFPSLPYSCSLAAAGLVL"	
Join(18574..18678,18821..18984,20752..21058,21870..22177,22377..22536)	
/gene="P0456E05.3"	
Join(18574..18678,18821..18984,20752..21058,21870..22177,22377..22536)	
/gene="P0456E05.3"	
/note="contains EST AU070785(R10238) similar to Arabidopsis thaliana chromosome 3, At3g48360 unknown protein"	
/codon_start=1	
/protein_id="BAB92604.1"	
/db_xref="GI:20804925"	
/translation="MCEGVRAAGDAAAADVITSSGRRRIPAHSTVLASAPVLES ILQRRLLKERDAAAGGKVRRAVLLIRGVTDAAAFLVRLIAGSSGDEEIDKSKAA QMLVAHAYRVPLKRCCEGASIGSLTAESVYVDTMQLAACDAPOLHURCTRLAKEF KAVEKTEAWRFLOENDPWLELDILQRLHDADLRKKRRKRAEGVYVLESEAMDCLS HICVTEGTVGVRGAPAAAPCPAYATACGLQLLIIRHFSRCHTSCPCRCMOLLIR LHAALCDLPDGHCTPTLCMQFRKEEKAATAKAKAGDDDDKGLLVKVRVARAMS SLGKRRQMSCQC"	
complement(join(23558..23638,24837..24941))	
/gene="P0456E05.4"	
complement(join(23558..23638,24837..24941))	
/gene="P0456E05.4"	
/note="contains EST AU162265(E30874) unknown protein"	
/codon_start=1	
/protein_id="BAB92605.1"	
/db_xref="GI:20804926"	
/translation="NVKVATTFAMTFGAFQWQMDRVHVHIALHQDEKKEMEREQE IKRMOAELMAQAKES"	
complement(join(26102..26407,26766..26901,27028..27170,27656..27784,28154..28402,28988..29359))	
/gene="P0456E05.5"	
complement(join(26102..26407,26766..26901,27028..27170,27656..27784,28154..28402,28988..29359))	
/gene="P0456E05.5"	
/note="hypothetical protein"	
magnesium transporter-like protein"	
/codon_start=1	
/protein_id="BAB92606.1"	
/db_xref="GI:20804927"	
/translation="MRPSAAGGGGGGRRKAAASRENLVVPASGARVEEA GKHAVMARTGLPARDLVLDPLISTEITLGERAIVNLERKAVITPAEVLNPSK DPAPASFVCDLOARVLASSDQAAEETDMGESAVTSFPALITTPPELEMTNKS NVVGGMTHSNMPTLTAADGNTKVLFPFEFRALEVCESACRSLEETSTLEQAYPA LDELTSKI STLNLNRYRQIKSLRVAISGRVKVRDEHLDDMDMAEMLYLEKIVR QEISTSSRVEVDDPSQLEVDREDEYRSEADYSNGTFIGYKPHIELEMLLEAYFVQI DGTNLKSLHREYDDTEDIINMLDDKQNLQMGVMLSTATVVTAGVAVVGLFGM NIGISLYADPTNEEKSRASMKFWETTLGTIAGCTVMYIVAGMGRKSILLO"	
Join(36977..37082,37255..37479,37607..37701,37823..38015,38424..38764,39169..39399,39493..39645)	
/gene="P0456E05.6"	
Join(36977..37082,37255..37479,37607..37701,37823..38015,38424..38764,39169..39399,39493..39645)	
/gene="P0456E05.6"	
/note="contains ESTs D46801(S11706),AU095929(S11706)"	
/codon_start=1	
/product="proline transport protein-like"	


```
QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16061 ATGGCTCACTACCGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGTCAACCCACC 16120
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValValValValValPheGlu 39
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16121 GCGCGCTCTCTCGCGCGCGCGTGCAGCAGCAGCAGTCTCAAGGAGAGCTTCCAG 16180
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisGly 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16181 GAGATCGAC-----AGTCCGCTCCGCCGCCGCCACACACACACACACCGGC 16231
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16232 AACGACTACTGTATGTCGCGAGACCAAGTGGAGGAGTCTCAACACCTGCACCGGC 16291
QY 80 GluValHisGluArgGluSerPheLeuAlaAtrGAlaAsn 93
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16292 GAGTCCGCGAGCGCAAGCAGAGCTCTCCGCTCAAGTCCGAC 16333

RESULT 3
AP003436/c 185095 bp DNA linear PLN 17-APR-2002
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION PAC clone:P0470A12.
ACCESSION AP003436
VERSION AP003436.2 GI:16197551
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
Clone:P0470A12
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
Clone:P0470A12
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 185095)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (21-MAR-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakienias@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/;
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT On Oct 17, 2001 this sequence version replaced gi:13430001.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLAST2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
Detailed information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/genomeSeq.html.
Location/Qualifiers
1. 185095
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"

FEATURES
source
```

```
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0470A12"
/join(10813..10841,11094..11205,11289..11432,11817..11921,
12046..12129,12315..12430,12535..12636,12715..12828,
12916..12972,13222..13288,13356..13409,13798..13872,
13956..13991,14389..14490)
/gene="P0470A12.1"
CDS
/join(10813..10841,11094..11205,11289..11432,11817..11921,
12046..12129,12315..12430,12535..12636,12715..12828,
12916..12972,13222..13288,13356..13409,13798..13872,
13956..13991,14389..14490)
/gene="P0470A12.1"
/note="hypothetical protein
similar to Oryza sativa chromosome 5, P0574H01.5"
/codon_start=1
/protein_id="BAB90276.1"
/db_xref="GI:20161352"
/translation="MSKLAILGRSSGLPRNEGMRLLFSAVIGVLMGLYGLFGISPTVN
VKLHFPSSIIISIEDKSDGITTOTLLNHWMTSANSKRNNSESDIPIKIVPTNP
KGAELAPGIYVPTDILYLRRLWGEPSDLTSQPRYLITFTVGSQKADAAVKKFS
ENFTIMLFHYDRTNDWDEFWSKRAIHVSRRQTKWYAKRFLPHLDIVAPYDIFILW
DEDLSVOHFNAEAVIKLVKHELSIQPLEPDKLGTQWMTKRLGDQEHKVTPEPG
WCTDPHLPCCAFAVEINATVFSRDAMRCVHMIONDLVHVGWGLDFALRRCVPEAHERKI
GVYDSQWVIHQVIFSLGNQGTAEGRTPWEGVRARCKEWGMFOKLADAERKAYILK
GITPPN"
complement(15118..16344)
/gene="P0470A12.2"
complement(15118..16344)
/gene="P0470A12.2"
/note="contains EST AU069076(C51993)
unknown protein"
/codon_start=1
/protein_id="BAB90277.1"
/db_xref="GI:20161353"
/translation="MAMARRNGGYGYGAYSYGYSPQVNYHSQSESVTVTVTKM
DEMSVHNAVOKMSSSFREKVEYEOQGAGGDVQVCRNGAAGAVORHTYEEKKEE
VYESNHGGGCGGCGGYAROHGAGGQKHAYOHEVLKGYESAGGYAAARHNGAGG
GVKQYSTRREEYDAGYARHNGGQVOKYSSYRRDEEDGAAAGGGRHGHAGGG
AVKHATYKHQAIEGVYCCGAGGYNHQAIVAGGSHGQHYGATAAAYGNASNKQ
HFTAAAGHHSGGHTYHQHSTEECEDDEDDDDDDDDDDDDDDDDDDDDDDDDDDDD
SVHSYQAYQHEERKAGGRNHAYERHEEHGGGGAQRYKIESSTQVYAGGGGNN
GCVVRGNKAPCFRINA"
complement(join(19564..19724,21382..21688,22059..22217,
22387..22505,22923..23035,23239..23318))
/gene="P0470A12.3"
complement(join(19564..19724,21382..21688,22059..22217,
22387..22505,22923..23035,23239..23318))
/gene="P0470A12.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90278.1"
/db_xref="GI:20161354"
/translation="MARRVGGTSEVGGGNGEARGRRRORSRRRARGARRRLPRRG
SPASPGRREAGGSGGGGRAPVINTTTRSRKPRWDRSGGRDHAASRHSIHT
RETNWNGAIKLRTFVGVREFRCCKPISLVTPINWAGILGCLVLRLLCLRQWR
RUAARHVAERNEALGVVVLRLHVPSPVAATMTRKQCATGAALPYEAHTIENTIGE
GEHEHDVROPQGLRTLRLHQOPRYESTSCARIGGSGRAGEKHTLNCQHRDRT
TDDQAVRMTTLKRKTKTCKSVCHGRVQVWQAPYQG"
complement(23929..24237)
/gene="P0470A12.4"
complement(23929..24237)
/gene="P0470A12.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90279.1"
/db_xref="GI:20161355"
/translation="MAQYYSVDHCAEMNRPPhAGGEHVAVRRESYEEVDEMARAGR
GHHNGGGGGHILGYSGSRSHDHALGGRHEHLVHGDEHRHGHGGGRQDSCGTQYIG"
complement(join(27239..27499,30542..30622,30720..30899,
30998..31165))
/gene="P0470A12.5"
```



```
/product="putative glycine-rich protein"
/protein_id="CAB61838.1"
/db_xref="GI:6478144"
/translation="MAHFKEMDYDEVTSMKPGRGRHGGGVQVQFVVKETFEVEQV
TPRGRSGHHGGHNRGNGHGHFQARETFERHQHARRVPRQAQGERPCQ"

BASE COUNT      143 a      153 c      154 g      112 t
ORIGIN

Alignment Scores:
Pred. No.:      3.68e-13      Length:      562
Score:          225.50      Matches:      59
Percent Similarity: 70.00%      Conservative: 11
Best Local Similarity: 59.00%      Mismatches:  20
Query Match:     44.22%      Indels:      11
DB:              8      Gaps:        6

US-10-090-035-2 (1-93) x SST242802 (1-562)
QY      1 MetAlaTyrTyrGlnGlu---ValAspTyrCysSerGluGluValArgSerValAlaPro 19
Db      63 ATGGCGCATTCAAGGAGATGGTGACTAC-----GAGGAGGTGACCTCGATGGCCAAA 116
QY      20 AlaGlyPheGlyArgHis---GlyGlyGlyValGlnGlnHisValValGlyGlyLysPhe 38
Db      117 CCGGCTTCGCGCCGCCCGCGCGCGCTCCAGCAGTTCGTCTGCTCAAGGAGACGCTTC 176
QY      39 GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHis 58
Db      177 GAGGAGGTTCGAGCAGTCCACCCCGCCGCGCAGTTCGCGCGCGCAGTTCGCGCGCAGC 233
QY      59 Gly-----GlyHisGly-----PheValValArgGluThrArgValGluGluAsp 73
Db      234 GCAACCAAGGCGCCAGCGCTCCGCGCCACTCCAGGCGCGCGAGACCAAGTTTGAGA--GAC 291
QY      74 IleAsnThrCysThrGlyGlyValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      292 ATCAACACGCGCAGCGCGAGTTCACGAGGCGAGGAGACGCTCTGTTCAGGCGCTGAC 351

RESULT 6
SME591784/c      300000 bp      DNA      linear      BCT 05-JUL-2002
LOCUS
DEFINITION      Sinorhizobium meliloti 1021 complete chromosome, segment 3/12.
ACCESSION      AL591784 AL591688
VERSION      AL591784.1 GI:15073438
KEYWORDS
SOURCE      Sinorhizobium meliloti.
ORGANISM      Sinorhizobium meliloti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.
REFERENCE      1 (bases 1 to 300000)
AUTHORS      Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J.,
Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreano, S., Gloux, S.,
Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D.,
Pohl, T., Portetelle, D., Puehler, A., Purnelle, B., Ramsperger, U.,
Renard, C., Thebault, P., Vandenbol, M., Weidner, S. and Galibert, F.
TITLE      Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)
MEDLINE      21396507
PUBMED      11481430
REFERENCE      2 (bases 1 to 300000)
AUTHORS      Gouzy, J.
TITLE      Direct Submission
JOURNAL      Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO
EU Consortium
COMMENT      MELILO EU Consortium:
Laboratoire de Biologie Molculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Genetique et Developpement UMR6061-CNRS,
Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,
D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,
```


Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte
des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,
B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@oulouse.inra.fr
http://sequence.toulouse.inra.fr/melliloti.html.

FEATURES

```

source
1..300000
/organism="Sinorhizobium melliloti"
/strain="1021"
/db_xref="taxon:382"
gene
115..1149
/genes="SMC02146"
115..1149
/genes="SMC02146"
/function="cell processes; transport of small molecules"
/notes="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE PHOSPHATE-BINDING PERIPLASMIC PROTEIN"
/protein_id="CAC45081.1"
/db_xref="GI:15073439"
/db_xref="SPTREMBL:Q92SA4"
/translation="MKSLLKLVAAALVSAAGFAGVAAADTOVAGSSTVLPYKIVAE
NFEFTPFKPTPVVSEGGSSAGLKFECKGVEDIDTANSRKINDSEVEACKAAGVT
ETQAEIGYDGVATDAGNADLALKPEHLKALAEIVVDGKLVANPYKWSSEVDAS
LPDIAIAYIGEKHGRVFEFEKILAAAGCKDSQAQDVKAASDERQKQSKCVARVK
DGMVADICDTETILARIASNKNGIVFGLSFYENNADKLKVAIVNGIVPSTETIASG
EYPSRPLFFVYKKAHIGVPLGEYVEFFVDQMDPDSPLAEYGLVAAPDAEREI
RTSFEAGNLL"
1241..1846
/genes="SMC02145"
1241..1846
/genes="SMC02145"
/function="miscellaneous; hypothetical/partial homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="HYPOTHETICAL SIGNAL PEPTIDE PROTEIN"
/protein_id="CAC45082.1"
/db_xref="GI:15073440"
/db_xref="SPTREMBL:Q92KQ0"
/translation="MSTSLILLIILIGFTAFIAGRARSPARSGGKLASLHSLPGYHG
SYVAIWAMLPALLILAVLWLIAPYIIDYTVRSLPEFPVQNGAATVQALGQVKSAA
GFERLSADELAALENGSTNNREVLASKGVALAGEGEPYMQSQAQFNAMROYSRWLMS
AVLLVSVAGGLYAIRNIDVFRRAAIGWSASCARSRRP"
1813..2724
/genes="pstC OR SMC02144"
1813..2724
/genes="pstC OR SMC02144"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : putative
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT SYSTEM PERMEASE ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45083.1"
/db_xref="GI:15073441"
/db_xref="SPTREMBL:Q92SA3"
/translation="MLGSLIVASSIAILTTVTVGIVASMLSEAGRFTTMSPPFFFTGV
WDPRFAAGSGAGSQFGLPLLAGTLYIAPVAMLEAVPICGLFAAIYMAEYASPRLS
TAKPLLEVLGIPITVIVGFFALVVGPFRLDISQLAGLVNQFNTQAGSVITAGI
VAGIMLPFVSSLSDDIITQVPRALRDGSLGATRSYTKVKIILPAALPGIIVGALL
TASRAGETMIVLVAAGVAARMQNLNPEMTTIVTKIVNQLTGDLEFTSPQTLVAFAL
GITLIVGLTGLNIYIVIRKYREQYE"
2721..4043
/genes="psta OR SMC02143"
2721..4043
/genes="psta OR SMC02143"
/function="cell processes; transport of small molecules"
/notes="Product confidence : putative
Gene name confidence : putative
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE PHOSPHATE TRANSPORT SYSTEM PERMEASE ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45084.1"
/db_xref="GI:15073442"
/db_xref="SPTREMBL:Q92SA2"
/translation="MTEFASPTVTVSTRPAAARRDIGIKRRYAAARRFRFAGMSATAI
GLFPLAVLMSVSKGYTAFQQTITLPIEFSEKVIDPKNERATDPKQLYTANYPLLV
REALVQLGLDPSKRLPVKAQADMVASARTLRDIWVADPSVIGRTIPVTLAEANI
DSAFKQIDLSVDETTRKVKDEQVAMNQLAEAGVLKSFNTGIFTFGASSRPERAGV
GVALGLSVMYMLIVLSLPIGVAAISYLEEPAPKRLTDLIEVINNLAAVPSIVYG
LIGLAVVNFAGPRAALGVFVLTMLPTIILITRAALKANVPSIIRKSLAAGLGS
KQMTIFHVLPLAMPGLITIGLHALETGAPLLIICWIAVAFVDFPQSPLEPSTAL
PVQIYMWANAEARAFVERISGAIIILLIIFLMNLGAILLRFRFRFR"
4057..4872
/genes="pstb OR SMC02142"
4057..4872
/genes="pstb OR SMC02142"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : probable
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT ATP-BINDING ABC
TRANSPORTER PROTEIN"
/protein_id="CAC45085.1"
/db_xref="GI:15073443"
/db_xref="SPTREMBL:Q92SA1"
/translation="NNINSEAAVEKALDQKNVTPYKMGNDVSVYGEKRALFDVNL
NIRENTVVALIGPSCGSGSTFRLNRRNDTIENCRVTGKITLDEDDIYDPSIDVEL
PARVGNVQKNPFPKSIYDNVSGPRILHARLKAELDEVETSLQKAGLWNEVKDR
LOEPQTLGSGGQORLCIARAVAVSPEVILMDEFCSDALPTATARKVELIHELRAFET
IVIVTHSMQQAARVSQRTAMFHLGNVEENDTKMFTNPDDQTDYIMGRFG"
4959..5672
/genes="phoU OR SMC02141"
4959..5672
/genes="phoU OR SMC02141"
/function="cell processes; transport of small molecules"
/notes="Product confidence : probable
Gene name confidence : probable
Predicted by Homology
Predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE PHOSPHATE TRANSPORT SYSTEM
TRANSCRIPTIONAL REGULATOR PROTEIN"
/protein_id="CAC45086.1"
/db_xref="GI:15073444"
/db_xref="SPTREMBL:Q92SA3"
/translation="MSHAHMSAPDEELKYLTRRISDMGGLBQMVADSVRALVNSDL

```

ALAOKVISDDTILDDAERQIGKAVITIAKROPMSDLREIMGSIRIAADLERVGDIG
KNTARVITAVAGSGIPKRLAGLEHLELALVQLKLELDVLTASPEKANSIRERDEE
IDAVYLSLFRRELLYMMEDPRNITPCTHLLFCARKNIERIGHATNIAETIYYMATGAQ
PQGRPKDDMTSLGVSVD"

gene

/gene="phob OR SMC02140"

CDS

/gene="phob OR SMC02140"

/function="small molecule metabolism; global functions;

global regulatory functions"

/note="predicted by Codon_usage

predicted by Homology

predicted by Framed"

/codon_start=1

/evidence-not_experimental

/transl_table=11

/product="PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY

PROTEIN"

/protein_id="CAC45087.1"

/db_xref="GI:15073445"

/db_xref="SWISS-PROT:Q52990"

Alignment Scores:

Pred. No.:	1,01e+03	Length:	300000
Score:	93.50	Matches:	28
Percent Similarity:	39.18%	Conservative:	10
Best Local Similarity:	28.87%	Mismatches:	28
Query Match:	18.33%	Indels:	31
DB:	1	Gaps:	5

US-10-090-035-2 (1-93) x SME591784 (1-300000)

Qy	15	ArgSerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValVal	34
Db	19988	CGTGATCTAAAGCCGAGGCG-----GGACAGCATGATGAAGTCGCGGAGACGACGAATG	19932
Qy	35	LysGluLysPheGlu-----GluValAspThrValSerArgAlaGly-----	48
Db	19931	CGTCATGCTTCGATGCGATGCTGCTGCGGTCATGAGTTCGCGGCGCGATGAG	19872
Qy	49	-----AlaAsnHisHisHisHisHisGlyHisGly-----	59
Db	19871	CGTCACGCGCAACCATCACCTCTGACGCGGCGACATGCGCATGATGAAGGCTGGCCCGT	19812
Qy	60	-----GlyHisGlyPheValValArgGlu	67
Db	19811	ATAAGGGCAATCCCTCCATGACGGGTACGGTCAGGATCACGGT-----CGAAAG	19761
Qy	68	ThrArgValGluGluAspIleAsnThrCysThrGlyGluValHisGluArg	84
Db	19760	CGAAGCGCGAGCGCGATGTTCCACCACTGCTGCTGCTGCGGCGCGC	19710

RESULT 7

AF003103	814 bp	mrna	linear	PLN 27-JUL-1997
LOCUS	Arabidopsis thaliana	AP2 domain containing protein RAP2.10	mrna,	
DEFINITION	partial cds.			

ACCESSION AF003103

VERSION AF003103.1

GI:2281644

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 814)

Okamoto, J.K., Caster, B., Villarroel, R., Van Montagu, M. and

Jofuku, K.D.

The AP2 domain of AP2AL2 defines a large new family of DNA

binding proteins in Arabidopsis

Proc. Natl. Acad. Sci. U.S.A. 94 (13), 7076-7081 (1997)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS Okamoto, J.K., Caster, B., Villarroel, R., Van Montagu, M. and

Jofuku, K.D.

TITLE Direct Submission

JOURNAL Submitted (08-MAY-1997) Biology, University of California, Santa

CRUZ, CA 95064, USA

FEATURES

source

1..814

Location/Qualifiers

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

misc_feature

1..814

/note="similar to EST with GenBank Accession Number

T76017"

CDS

<1..781

/note="putative DNA binding protein; similar to A.

thaliana AP2AL2 encoded by GenBank Accession Number

U12546"

misc_feature

278..445

/note="encodes AP2 domain"

BASE COUNT

218 a 151 c 233 g 212 t

ORIGIN

Alignment Scores:

Pred. No.:	2.82	Length:	814
Score:	92.50	Matches:	24
Percent Similarity:	50.70%	Conservative:	12
Best Local Similarity:	33.80%	Mismatches:	24
Query Match:	18.14%	Indels:	11
DB:	8	Gaps:	5

US-10-090-035-2 (1-93) x AF003103 (1-814)

Qy	21	GlyPheGlyArgHisGlyGlyGly---ValGlnGlnHisValValLysGluLysPheGlu	39
Db	485	GCAGGAGGAGTCAACGGTGTGAGATATGTCGCGGCGGTATATAAGGAGAAACGGCG	544
Qy	40	Glu-----ValAspThrValSerArgAlaGlyAla-----AsnHisHis	52
Db	545	GAGGTGGACACAAAGTGGATCGGTAGAACGCGGCGGCGGGAATCGTCATCAT	604
Qy	53	HisHisGlyHisHisGlyHisGlyPheValValArgGluThr-----Arg	69
Db	605	CATCATCATCAACATCAACGTCGTAATCATGATTACGTAGATAGTATAGTATTATCGT	664
Qy	70	ValGluGluAspIleAsnThrCysThrGlyGlu	80
Db	665	ATTATGATGATGCTTATGAGTGTAGTAGTAA	697

RESULT 8

ATAJ2598

LOCUS

Arabidopsis thaliana

DEFINITION

AJ002598

ACCESSION

VERSION

AJ002598.1

GI:2632062

KEYWORDS

TINY-like protein.

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 815)

Terry, N., Heijnen, L., De Keyser, A., Van Asseldonck, M., De

Clercq, R., Verbakel, H., Gielen, J., Zabeau, M., Villarroel, R.,

Jesse, T., Neyt, P., Hogers, R., Van Den Daele, H., Ardiles, W.,

interuniversity institute for biotechnology, K.L. Ledeganckstraat
35, 9000 Gent, BELGIUM

COMMENT
this cDNA is corresponding to EST 140D13 (T76017).

2. (bases 1 to 1009)
 Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Banh, J.,
 Bowser, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y.,
 Ishida, J., Jtlang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M.,
 Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K.,
 Quach, H. L., Sakurai, T., Sato, M., Seki, M., Southwick, A., Tang, C. C.,
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
 Davis, R. W., Theologis, A. and Ecker, J. R.

Toriiumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (24-Sep-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CDNAs (RAFL cDNA : "RIKEN Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCPEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Xia, C.J., Koeseema, E., Meyers, M. C., Shinn, P., Banh, J., Bowser, L., Dale, J. M., Goldsmith, A. D., Jiang, P. X., Jones, T., Karlin-Neumann, G., Li, M., Lee, J. M., Lin, J., Liu, X. X., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Southwick, A., Toriumi, M., Yanada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. C., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

```

Location/Qualifiers
1. .1009
   /organism="Arabidopsis thaliana"
   /db_xref="taxon:3702"
   /chromosome="4"
   /clone="RAFL09-50-B19(R14720)"
   /note="ecotype: Columbia"
1. .201
202. .792

```

```
res:          3.53      Length:    1009
266 a   183 c   269 g   291 t
793..1009
/replaces="g"
/note="compared to genome"
reference 433
DDLMFCSKSGEGFKRCNGSLERVDLNLPPETSDDD"
NGGDMSSAAY IRRKAAEVAQAQVDALAEAGAGGNRHHHQQHQRGNHHDYVDNHSIDRYIN
```

```
res: 3.53 Length: 1009
```


exon 12203..12795
/gene="AT4g36750"
/number=1
intron 12796..13547
/gene="AT4g36750"
/number=1
exon 13548..13776
/gene="AT4g36750"
/number=2
gene 14283..17541
/gene="AT4g36760"
join(14283..14342,14482..14545,14645..14784,14874..15158,
15237..15632,15717..15803,15889..15953,16062..16128,
16205..16339,16440..16520,16618..16672,16768..16872,
17002..17073,17162..17286,17374..17541)
/gene="AT4g36760"
/note="similarity to cytoplasmic aminopeptidase P, Rattus
norvegicus, PAX:G2760920
contains EST gb:N96008, AI998300.1"
/codon_start=1
/product="aminopeptidase-like protein"
/protein_id="CAB80342.1"
/db_xref="GI:7270625"
/translation="MASHPPLDALVVPSEYHQSEYVSARDKRRFEYVSGSGSAGLA
LITKRLRWTDGRYFLQALQQLSDENILMRGDEPLVEVWMSDNLPEANIGVDSWC
VSYDTANRWKGFARKNQKLITTTDLDEVWKRSPPEMSPVYVHPLEFAGRSYSHK
FEDLRALKQEGARGVIAALDEVAWLINRGTDAVCPVWHAFALITDSFLYDK
KVDSEANSGVFNVEYREYTDVISDALLASDRILISFASKTVOHEAAKMEIDSD
QDRLVWDNASCYALYKGLDAEKVLLQSPISLSKALKNPVELEGINKAHVRDGAAG
VOYLWLDNMOELYAGSGYFLEAEAKKPKSETSKLTVTVSDKLESLRASKHEFRG
LSPTSSVGSNAVTHYSPEEACAEKDPKTYLCDGAOYLDGTTDITRVVHFGKP
SAHEKCYTAVFGHVALGNARFPKGTNGYTDILARAPLMKYGLDYRGTGCVGSY
LCVHEGPHQVSRPARNPLQATMTVDPEGYEDCGNIGRLENVLVNDATETENF
KRGYIQLQFERITWAPYQVQLDIDELTREIDWLNTYHSKCKDILAPFNMQTEMELK
KATEPVSVSA"
exon 14283..14342
/gene="AT4g36760"
/number=1
intron 14343..14481
/gene="AT4g36760"
/number=1
exon 14482..14545
/gene="AT4g36760"
/number=2
intron 14546..14644
/gene="AT4g36760"
/number=2
exon 14645..14784
/gene="AT4g36760"
/number=3
intron 14785..14873
/gene="AT4g36760"
/number=3
exon 14874..15158
/gene="AT4g36760"
/number=4
intron 15159..15236
/gene="AT4g36760"
/number=4
exon 15237..15632
/gene="AT4g36760"
/number=5
intron 15633..15716
/gene="AT4g36760"
/number=5
exon 15717..15803
/gene="AT4g36760"
/number=6
intron 15804..15888
/gene="AT4g36760"
/number=6
exon 15889..15953
/gene="AT4g36760"
/number=7
intron 15954..16061
/gene="AT4g36760"
/number=7
exon 16062..16128
/gene="AT4g36760"
/number=8
intron 16129..16204
/gene="AT4g36760"
/number=8
exon 16205..16339
/gene="AT4g36760"
/number=9
intron 16340..16439
/gene="AT4g36760"
/number=9
exon 16440..16520
/gene="AT4g36760"
/number=10
intron 16521..16617
/gene="AT4g36760"
/number=10
exon 16618..16672
/gene="AT4g36760"
/number=11
intron 16673..16767
/gene="AT4g36760"
/number=11
exon 16768..16872
/gene="AT4g36760"
/number=12
intron 16873..17001
/gene="AT4g36760"
/number=12
exon 17002..17073
/gene="AT4g36760"
/number=13
intron 17074..17161
/gene="AT4g36760"
/number=13
exon 17162..17286
/gene="AT4g36760"
/number=14
intron 17287..17373
/gene="AT4g36760"
/number=14
exon 17374..17541
/gene="AT4g36760"
/number=15
gene 17779..19152
/gene="AT4g36770"
complement(17779..19152)
/gene="AT4g36770"
complement(17779..19152)
/gene="AT4g36770"
complement(17779..19152)
CDS
/note="strong similarity to UDP-glucose
Glucosyltransferase, Manihot esculenta, PIR2:S41951
Contains UDP-glucosyltransferases signature AA343-386
contains EST gb:AI998300.1"
/codon_start=1
/product="glucosyltransferase-like protein"
/protein_id="CAB80343.1"
/db_xref="GI:7270626"
/translation="MELHGVASPGMGHAPVILELGRHLLNHHGDRVTVFLVTDV
SRSKSLIGKTEEDPKFVIRFIPDVSGDLSGLTLKLAEMMRKALPEIKSSVMEL
EPRPVFVVDLLGLEALEVAKELGIMRKHVLTTSAMFLATVYMASLDKQELYKOLS
SIGALLIPGSPVKFERAQDPKRYIRELAESQRIQIGDEVITADGVFNTWHSLEQVFIG
SFLDPENLGRVMRGVPVYPVGLVRPAEPGLKHGVLWDLDLPKESVYVLLGVVGCAL
TFQETNELAYGLELTHGRFVWVVRPAEDDPSASMEFDKTKNETEPDLDFLPNGFLDRK
DIGLVRTWAPQEEILAHKSTGGVFTCHGWSVLESIVNGVPMPLYSQKMNARM
VSGELKTAQINVADGIVKKEVIAENVKRVMDDEEGKMKRKNVKKELKTAEEALNMTH
IPSAVPT"
complement(17779..19152)
exon

```

/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="BAC TAMU8H13"
68418. .87492
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.5E03"
87372. .105097
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.4G12"
103943. .129912
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.6B05"
126229. .145429
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.3F11"
144619. .166550
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.4F04"
166717. .183138
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.3A06"
189687. .192093
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.4B06"
192023. .206420
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="cosmid C1C7A10.4E12"
3. .560
/gene="C7A10.10"
<3. .560
/gene="C7A10.10"
/!note="similarity to cytochrome P450, Nicotiana tabacum,
P4RX-E118611
Contains Cytochrome P450 cysteine heme-iron ligand
signature [FGLGRRACPG]"
/codon_start=1
/product="cytochrome like protein"
/protein_id="CAB16768.1"
/db_xref="GI:4006850"
/translation="SLLNANPEVLNKARDEIRMGIVDRLLSESDIPNLYLQNIYVSET
LRKYPAAPMLLPVHACKVCKGVGYDMPRGTMLLTNAWIHRDPLLWDDPTFKPKPREF
KEGEAKLMLPGFLGRRACPGSGSLAQRVLTSLGSLIQCFEWERIGEEBVDNTEGPGLT
MPKARPLEAMCARDEYFVGKILPDSS"
3. .560
/gene="C7A10.10"
/number=1
/number=1
978. .2731
/gene="C7A10.20"
Join(978. .1475,1630. .2016,2114. .2731)
/gene="C7A10.20"
/!note="similarity to cytochrome P450 Glycyrhiza echinata,
PATY:D10232987
Contains Cytochrome P450 cysteine heme-iron ligand
signatures [FGLGRRACPG] and 474:[FGLGRRACPG]"
/codon_start=1
/product="cytochrome P450-like protein"
/protein_id="CAB16769.1"
/db_xref="GI:4006851"
/db_xref="SPTREMBL:O23154"
/translation="MEQQTILFFLSLTFTIGRIKRRPLPPSPWALPVIGHL
RLKLPVHRVFLSVSESGDAPISLGRNLFVYVSSHSLAECEFTKNDVLIANRFN

```

```

SLASHISGIVCTVVTASYGDHNRNLRIGAVEIFSAHRLNSFSIRDEIHLIACI
SNSSLETKVMKMFNLTNNIIRMLAGKCYGDAEDDPAKRVRELIABGMGC
FAGANTADYPLTWITGSKRIKKIASRLDEFQGLVDERREGKRRQNTMVDHLLC
LQETPEYTNLIKIMLSILACTDTSVLTWTLISALLNHPOLLSKARDEIDNKV
GNRQVSESDLSHLYLQNTVSESLRYPASPLVPHVASEDCVGGVHMPRGTMILT
NAWAHRDPIKDDPTSFPERFEKEGEAKLLGLGRRACPGSLAORLASLTIGS
LIQCFEWEIRIGEEDVMTTEGGGVIMPKAIPLVAMKARPVVGKILNESA"
978..1475
/feature="C7A10.20"
/number=1
1476..1629
/feature="C7A10.20"
/number=1
1630..2016
/feature="C7A10.20"
/number=2
2017..2113
/feature="C7A10.20"
/number=2
2114..2731
/feature="C7A10.20"
/number=3
3115..5137
/feature="C7A10.30"
Join(3115..3615,4090..4464,4535..5137)
/feature="C7A10.30"
/feature="strong similarity to cytochrome P450, Glycyrrhiza
echinata, PATX:D1023287
Contains Cytochrome P450 cysteine heme-iron ligand
signature 431:[FGMGRACPG]
contains EST gb:T43640, Aa395149, T42716, T41670"
/codon_start=1
/product="cytochrome P450-like protein"
/protein_id="CAB16753.1"
/db_xref="GI:2464850"
/db_xref="SPTREMBL:O23155"
/translacion="MDLTQILLSLFLTISIKLLTKGNRKNLPSPAYPLPVIGH
LHLKOPVHTFHSISKSGNAPVFLHRLGNRLVYVYSSIAEECTFNDVVLNRP
KFTTSKHLGNATVLLSASGDHNRNLRRAAIVEIFSHRLNSIRKDEIRRLISH
LSRSLHGFEVEMKTLTINLASNTTIRMLAGKRYFGEDNDKALVKNLYSEAVTSAG
AGNFDILSLRWVSSYEKRIKLNEDFTLOKLVDKRAEKEGETMIDHLLAQD
TOPDYDVIIRKILILITAGTDTSSVTLWAMSNLLNHPILKARMEIDKVGLD
RVDESIDVNLISQISIVLETIRMPVAPVLLPHLSDECKVGGYDIPSGTMYLTNAW
AMHRDPEVWEDPEIFKPERFEKEGEAKLLISFGMGRACPGAGLAHLINQALGSLVQ
CFEWEVGEDVDMTEKDGATLPKAIPLAMKARSIVDKLI"
3115..3615
/feature="C7A10.30"
/number=1
3616..4089
/feature="C7A10.30"
/number=1
4090..4464
/feature="C7A10.30"
/number=2
4465..4534
/feature="C7A10.30"
/number=2
4535..5137
/feature="C7A10.30"
/number=3
5994..7942
/feature="C7A10.40"
Join(5994..6494,6880..7263,7340..7942)
/feature="C7A10.40"
/feature="similarity to cytochrome P450 monooxygenase,
Arabidopsis thaliana, PATCH:D1029478
Contains ATP/GTP-binding site motif A (P-loop) [ARAIVGKT],
Cytochrome P450 cysteine heme-iron ligand signature
[FGMGRACPG]"
/codon_start=1
/product="cytochrome P450-like protein"
/protein_id="CAB16770.1"
/db_xref="GI:4376088"

```

```

/db_xref="SPTREMBL:O23156"
/translacion="MDLTQILLSLFLTISIKLLTKGNRKNLPSPAYPLPVIGH
LHLKOPVHTFHSISKSGNAPVFLHRLGNRLVYVYSSIAEECTFNDVVLNRP
DIIMKGVNFTNMAASVGDHNRNLRRAAIVEIFSHRLNSIRKDEIRRLITH
LSRSLHGFEVEMKTLTINLASNTTIRMLAGKRYFGEDNDKALVKNLYSEAVTSAG
AGNFDILSLRWVSSYEKRIKLNEDFTLOKLVDKRAEKEGETMIDHLLAQD
TOPDYDVIIRKILILITAGTDTSSVTLWAMSNLLNHPILKARMEIDKVGLD
RVDESIDVNLISQISIVLETIRMPVAPVLLPHLSDECKVGGYDIPSGTMYLTNAW
AMHRDPEVWEDPEIFKPERFEKEGEAKLLISFGMGRACPGAGLAHLINQALGSLVQ
LIQCFEWEVGEAEVDMTEGEGITMPKATPLRAMCKARAIVGKTI"
5994..6494
/feature="C7A10.40"
/number=1
6495..6879
/feature="C7A10.40"
/number=1
6880..7263
/feature="C7A10.40"

Alignment Scores:
Pred. No.: 855 Length: 206420
Score: 92.50 Matches: 24
Percent Similarity: 50.70% Conservative: 12
Best Local Similarity: 33.80% Mismatches: 24
Query Match: 18.14% Indels: 11
DB: 8 Gaps: 5

US-10-090-035-2 (1-93) x ATAP21 (1-206420)
QY 21 GlyPheGlyArgHisGlyGlyGly---ValGlnGlnHisValValLysGlyLysPheGlu 39
||| ||| :||:||||||| :||: |||: |||
Db 178372 GGAGGAGGAGTGAACGGTGGTGGAGATATGTCGGCGCGGTATATAAGGAGAAAGCGCGC 178313
QY 40 Glu-----ValAspThrValSerArgAlaGlyAla-----AsnHisHis 52
||| ||| :||:||||||| :||: |||: |||
Db 178312 GAGGTTGGAGCAACAAGTCGTCGTAGAACGGCGGGGGAGGAGAAATCGTCATCAT 178253
QY 53 HisHisHisGlyHis---HisGlyGlyHisGlyPheValValArgGluThr-----Arg 69
||||||| ||| ||| ||| ||| ||| ||| |||
Db 178252 CATCATCATCAACATCAACGCTGTAATCATGATTAGTAGATAGTATAGTATGATTATCGT 178193
QY 70 ValGluGluAspIleAsnThrCysThrGlyGlu 80
||| :||:||||| :||: |||: |||
Db 178192 ATTAATGATGATGCTTATGAGTGTAGTAGTAA 178160

RESULT 13
MUSDELTA 1898 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse delta transcription factor mRNA, complete cds.
DEFINITION M74590
ACCESSION M74590
VERSION M74590.1 GI:192940
KEYWORDS delta-transcription factor.
SOURCE Mus musculus (strain C57BL/6J, sub_species domesticus) cDNA to
mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1898)
REFERENCE Hariharan,N., Kelley,D.E. and Perry,R.P.
AUTHORS Delta, a transcription factor that binds to downstream elements in
TITLE several polymerase II promoters, is a functionally versatile zinc
finger protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (21), 9799-9803 (1991)
MEDLINE 92052178
PUBMED 1946404
FEATURES
source
Location/Qualifiers
1..1898
/organism="Mus musculus"
/strain="C57BL/6J"
/sub_species="domesticus"
/db_xref="taxon:10090"
43..1287
/standard_name="delta"
CDS

```



```

/codon_start=1
/product="delta-transcription factor"
/protein_id="AAA37521.1"
/db_xref="GI:192941"
/translating="MAGSDFLYATDGSSEPAEIVELHEIETVPIETTVVVEE
EEEDDDGGGHHGGGGHGHGHHHHHHHHHPMIALQPLVTDDPTQVHHH
QEVLVOTREVVGGDSGLRAEDFEOILIPVPAPAGGDDDIETQLVTVAAGK
SGGASGGGRVKKGKKSKGKSYLGGAGAAGGADPGNKKWQKVQIKTLEGE
SSVTHMSDEKDKIDHTVVEEQIIGENSPDYSEYMTGKLPPGIGIDLSPKOL
AEFAPRKI KEDDAPTACPHKCTKMRDNAMRKHLLHTHGPRVHVCAEGKAF
VESKLRHLQVHTGKEPFCQTEGCGKRFSLDENLRTVHVIHTGDRPVVCFPCGCK
KFAOSTNLKSHILTAKANNQ"
BASE COUNT      509 a   453 c   525 g   411 t
ORIGIN
Alignment Scores:
Pred. No.:       7.56          Length:      1898
Score:           92.00         Matches:     22
Percent Similarity: 40.85%    Conservative: 7
Best Local Similarity: 30.99% Mismatches:   26
Query Match:      18.04%     Indels:      16
DB:               10        Gaps:         2

US-10-090-035-2 (1-93) x MUSDELTA (1-1898)
Qy 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 205 GGCGGGCCGCCACCACCGCGCGGGCGGCCAC----- 240
Qy 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGlyGly 60
Db 241 -----GGGCACCGCGCCACCACCATCACCACCCACCACCCACCACCCAC 288
Qy 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
Db 289 CCSCCATGATCGTCGTGACGCGCGTGTTGACGGAGCAC-----CCGACCCAA 336
Qy 81 ValHisGluArgArgGluSerPheLeuAlaArg 91
Db 337 GTGCACACACACAGGAGGTGATCTGTGTGCAG 369

RESULT 14
MUSUCRB
LOCUS      MUSUCRB      2330 bp      mRNA      linear      ROD 27-APR-1993
DEFINITION Mus musculus UCR-motif DNA-binding protein (UCRB) mRNA, complete cds.
ACCESSION M73963.1 GI:202270
VERSION    1 (bases 1 to 2330)
KEYWORDS   DNA-binding protein; zinc-finger protein.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  Flanagan,J.R., Becker,K.G., Ennist,D.L., Gleason,S.L., Driggers,P.H., Levi,B.Z., Appella,E. and Ozato,K. Cloning of a negative transcription factor that binds to the upstream conserved region of Moloney murine leukemia virus Mol. Cell. Biol. 12 (1), 38-44 (1992) 92107191 PUBLISHED 1309593
FEATURES   Location/Qualifiers
            1..2330
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /cell_line="F9EC tk-"
             /cell_type="teratocarcinoma"
             1..2330
              /gene="UCRB"
              1..102
              /gene="UCRB"
              103..1347
              /gene="UCRB"

/codon_start=1
/function="MuLV UCR-motif DNA-binding protein"
/codon_start=1
/evidence=experimental
/product="UCR-motif DNA binding protein"
/protein_id="AAA40522.1"
/db_xref="GI:202271"
/translating="MASGDLYIATDSEMPAIEVELHEIETVPIETTVVVEE
EEEDDDGGGHHGGGGHGHGHHHHHHHHHPMIALQPLVTDDPTQVHHH
QEVLVOTREVVGGDSGLRAEDFEOILIPVPAPAGGDDDIETQLVTVAAGK
SGGASGGGRVKKGKKSKGKSYLGGAGAAGGADPGNKKWQKVQIKTLEGE
SSVTHMSDEKDKIDHTVVEEQIIGENSPDYSEYMTGKLPPGIGIDLSPKOL
AEFAPRKI KEDDAPTACPHKCTKMRDNAMRKHLLHTHGPRVHVCAEGKAF
VESKLRHLQVHTGKEPFCQTEGCGKRFSLDENLRTVHVIHTGDRPVVCFPCGCK
KFAOSTNLKSHILTAKANNQ"
BASE COUNT      635 a   552 c   601 g   542 t
ORIGIN
Alignment Scores:
Pred. No.:       9.34          Length:      2330
Score:           92.00         Matches:     22
Percent Similarity: 40.85%    Conservative: 7
Best Local Similarity: 30.99% Mismatches:   26
Query Match:      18.04%     Indels:      16
DB:               10        Gaps:         2

US-10-090-035-2 (1-93) x MUSUCRB (1-2330)
Qy 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 265 GGCGGGCCGCCACCACCGCGCGGGCGGCCAC----- 300
Qy 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGlyGly 60
Db 301 -----GGGCACCGCGGCCACCACCATCACCACCCACCACCCACCACCCAC 348
Qy 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
Db 349 CCGCCCATGATCGCGTGTGACGCGGTGTGACGGAGCAC-----CCGACCCAA 396
Qy 81 ValHisGluArgArgGluSerPheLeuAlaArg 91
Db 397 GTGCACACACACAGGAGGTGATCTGTGTGCAG 429

RESULT 15
MUSTRANSO1
LOCUS      MUSTRANSO1      3041 bp      DNA      linear      ROD 27-JUL-1993
DEFINITION Mouse delta/Yyl/NF-E1/UCRBP transcription factor, exon 1.
ACCESSION L13969
VERSION    1 GI:293843
KEYWORDS   Transcription factor; zinc-finger protein.
SEGMENT    1 of 5
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  Safrany,G. and Perry,R.P. Characterization of the mouse gene that encodes the delta/Yyl/NF-E1/UCRBP transcription factor Proc. Natl. Acad. Sci. U.S.A. 90 (12), 5559-5563 (1993) 93296177 PUBLISHED 8516301
FEATURES   Location/Qualifiers
            1..3041
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /tissue_type="liver"
             /dev_stage="4-6 weeks"
             /tissue_lib="129SV Stratagene"

```

```
promoter      1076.  .1808
protein_bind 1728. .1737
              /bound_molety="SP1"
exon         1776. .2885
              /number=1
```

BASE COUNT	676 a	886 c	980 g	499 t
ORIGIN				

Alignment Scores:		
Pred. NO.:	12.3	3041
Score:	92.00	
Percent Similarity:	40.85%	
Best Local Similarity:	30.99%	
Query Match:	18.04%	
DR:	10	2
	Length:	3041
	Matches:	22
	Conservative:	7
	Mismatches:	26
	Indels:	16
	Gaps:	2

US-10-090-035-2 (1-93) x MUSTRANS01 (1-3041)

QY	21	GlyPheGlyArgHisGlyGlyValAlaGlnHisValLysGluLysPheGluGlu	40
Dd	2369	GGCGGGCGGCACCGCGCGGGCGGCCAC-----	2404
QY	41	ValAspThrValSerArgAlaGlyAlaAsnHisHisHisGlyHisHisGlyGly	60
Dd	2405	-----GGGCAGCGCGGCCACCACCAATCACCCACACCCACCAACCACC	2452
QY	61	HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu	80
Dd	2453	CCGCCCATGATCGCTGCAGCGCGTGGTGACGGACGAC-----CCGACCCCA	2500
QY	81	ValHisGluArgArgGluSerPheLeuAlaArg	91
Dd	2501	GTGCACCAACCACGAGGAGTGATCCTGGTCGAC	2533

Search completed: June 16, 2003, 11:37:14
Job time : 2119.5 secs

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 496)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.

FEATURES

source

1..496
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT
ORIGIN

97 a 156 c 153 g 90 t

Query Match

Best Local Similarity 67.5%; Score 456.4; DB 10; Length 496;
Matches 471; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 67 GACACCAAGCGTCTGCACCAATGGCTTACTACAGAGGTGGACTTACTGCTCGGAGGA 126
DB 7 GACACCAAGCGTCTGCACCAATGGCTTACTACAGAGGTGGACTTACTGCTCGGAGGA 66
QY 127 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGACGCGGCGGCGGTCCAGCAGCAGT 186
DB 67 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGACGCGGCGGCGGTCCAGCAGCAGT 126
QY 187 CGTCAAGGAGAGTTCAGGAGGTGCGACGGTATCACGCGCGCGGCCAA-CCACCACC 245
DB 127 CGTCAAGGAGAGTTCAGGAGGTGCGACGGTATCACGCGCGCGGCCAA-CCACCACC 186
QY 246 ACCACCATGTCACACGCGCGGCGGCTTCGTTGCGGAGACCGAGGTCGAGGAGG 305
DB 187 ACCACCATGTCACACGCGCGGCGGCTTCGTTGCGGAGACCGAGGTCGAGGAGG 246
QY 306 ACATCAACACCTCGACGCGGAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTA 365
DB 247 ACATCAACACCTCGACGCGGAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTA 306
QY 366 ACTGAGCGCGCGCGGCGGATCCACGCGGTCGCTGCTGGGTGCTTATGTA 425
DB 307 ACTGAGCGCGCGCGGCGGATCCACGCGGTCGCTGCTGGGTGCTTATGTA 366
QY 426 TGTCTGTGTTGACTGGTGTGTGAGGGTATCGTACTTGGCTATCGTACGTGCGGCACT 485
DB 367 TGTCTGTGTTGACTGGTGTGTGAGGGTATCGTACTTGGCTATCGTACGTGCGGCACT 426
QY 486 CAGCTCCTGTAGCAATAGCAATAAGCTCGTACCTGAATAAACTTCTTCGTAAT 543
DB 427 CAGCTCCTGTAGCAATAGCAATAAGCTCGTACCTGAATAAACTTCTTCGTAAT 484

Search completed: June 16, 2003, 03:55:49

Job time : 1458.61 secs

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.
Location/Qualifiers
1. 539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603" - stressed root cDNA library from
Wang/Bohnert lab
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+) XR;
Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 91 a 157 c 172 g 119 t

ORIGIN

Query Match 58.0%; Score 460; DB 9; Length 539;
Best Local Similarity 97.6%; Pred. No. 2.9e-41;
Matches 488; Conservative 0; Mismatches 8; Indels 4; Gaps 2;

QY 67 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGGAGGA 126
DB 499 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGGAGGA 440

QY 127 GGTGAGTCTGGTGGCCCGCGGCTTCGGCGCCACACGGCGGCGGCTCCAGCAGCAGCT 186
DB 439 GGTGAGTCTGGTGGCCCGCGGCTTCGGCGCCACACGGCGGCGGCTCCAGCAGCAGCT 380

QY 187 CGTCAAGGAGAGTTCGAGGAGTTCGACAGCGTATCACGGCGCGGCGGCA---ACCACCA 243
DB 379 CGTCAAGGAGAGTTCGAGGAGTTCGACAGCGTATCACGGCGCGGCGGCA---ACCACCA 320

QY 244 CCACACCATGTCACACACGGCGGCGGCTTCGGTGGTGGCGGAGGAGTTCGCTCGGAGGA 303
DB 319 CCACACCATGTCACACACGGCGGCGGCTTCGGTGGTGGCGGAGGAGTTCGCTCGGAGGA 260

QY 304 GGACATCAACACCTGACACCGCGGAGGTCCACAGCGGAGGAGTTCCTCCCGAGGCG 363
DB 259 GGACATCAACACCTGACACCGCGGAGGTCCACAGCGGAGGAGTTCCTCCCGAGGCG 200

QY 364 TAATGAGCGCGCGCGGCGGCGGCTTCACCGCGGCTTCGCTGGCTGGCTGGCTGGCTATG 423
DB 199 TAATGAGCGCGCGCGGCGGCGGCTTCACCGCGGCTTCGCTGGCTGGCTGGCTGGCTATG 140

QY 424 TATGCTGTGGTGTGACTGGTGTGTCAGGAGTTCATGCTGCTGGCTGGCTGGCTGGCTATG 483
DB 139 TATGCTGTGGTGTGACTGGTGTGTCAGGAGTTCATGCTGCTGGCTGGCTGGCTGGCTATG 80

QY 484 CTCAGCTCTCTAGCAATAGCAATAAGCTGACCTGAATAAACTTCTTCGTAAT 543
DB 79 CTCAGCTCTCTAGCAATAGCAATAAGCTGACCTGAATAAACTTCTTCGTAAT 21

QY 544 ACTAATACCTACATCAAAA 563
DB 20 ACTAATACCTACATCAAAA 1

RESULT 15
AW787314
LOCUS
DEFINITION
945002E06.X1 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.
ACCESSION
AW787314
VERSION
AW787314.1 GI:7844111
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays

AACTGGAGAAATTCGGCGCGCAGGAAATTTTTTTTTTTTTTTTTT. The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT73PAC vector. The library
then went through one round of normalization to Cot value
of 5 based on the methods of Marcello Bento Soares (Genome
Research 6: 791-806, 1996).
Research 6: 791-806, 1996).
1 others

BASE COUNT 124 a 174 c 159 g 92 t

ORIGIN

Query Match 68.5%; Score 462.8; DB 12; Length 550;
Best Local Similarity 96.8%; Pred. No. 1.4e-41;
Matches 482; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

QY 67 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGGAGGA 126
DB 53 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGAGAGTGGACTACTGCTCGGAGGA 112

QY 127 GGTGAGTCTGGTGGCCCGCGGCTTCGGCGCCACACGGCGGCGGCTCCAGCAGCAGCT 186
DB 113 GGTGAGTCTGGTGGCCCGCGGCTTCGGCGCCACACGGCGGCGGCTCCAGCAGCAGCT 172

QY 187 CGTCAAGGAGAGTTCGAGGAGTTCGACAGCGTATCACGGCGCGGCGGCA---ACCACCA 243
DB 173 CGTCAAGGAGAGTTCGAGGAGTTCGACAGCGTATCACGGCGCGGCGGCA---ACCACCA 232

QY 244 CCACACCATGTCACACACGGCGGCGGCTTCGGTGGTGGCGGAGGAGTTCGCTCGGAGGA 303
DB 233' CCACACCATGTCACACACGGCGGCGGCTTCGGTGGTGGCGGAGGAGGAGTTCGCTCGGAGGA 292

QY 304 GGACATCAACACCTGACACCGCGGAGGTCCACAGCGGAGGAGTTCCTCCCGAGGCG 363
DB 293 GGACATCAACACCTGACACCGCGGAGGTCCACAGCGGAGGAGTTCCTCCCGAGGCG 352

QY 364 TAATGAGCGCGCGCGGCGGCGGCTTCGGTGGTGGCGGAGGAGTTCGCTCGGAGGA 423
DB 353 TAATGAGCGCGCGCGGCGGCGGCTTCGGTGGTGGCGGAGGAGTTCGCTCGGAGGA 412

QY 424 TATGCTGTGGTGTGACTGGTGTGTCAGGAGTTCATGCTGCTGGCTGGCTGGCTGGCTATG 483
DB 413 TATGCTGTGGTGTGACTGGTGTGTCAGGAGTTCATGCTGCTGGCTGGCTGGCTGGCTATG 472

QY 484 CTCAGCTCTCTAGCAATAGCAATAAGCTGACCTGAATAAACTTCTTCGTAAT 543
DB 473 CTCAGCTCTCTAGCAATAGCAATAAGCTGACCTGAATAAACTTCTTCGTAAT 532

QY 544 ACTAATACCTACATCAAAA 561
DB 533 CCTTAATAAAATCAAAA 550

RESULT 14
AI855425/c
LOCUS
DEFINITION
603016F02.X1 603 - stressed root cDNA library from Wang/Bohnert lab
Zea mays cDNA, mRNA sequence..
ACCESSION
AI855425
VERSION
AI855425.1 GI:5499558
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 539)
Maize ESTs from various cDNA libraries sequenced at Stanford
Walbot, V.
University
Unpublished (1999)
JOURNAL
COMMENT
Unpublished (1999)
Walbot V

clade; Panicoideae; Andropogoneae; Zea.						
REFERENCE	1	(bases 1 to 553)				
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingrey,S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2	(bases 1 to 553)				
AUTHORS	Coe,E.C.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
FEATURES	Location/Qualifiers					
source	1..553					
	/organism="Zea mays"					
	/db_xref="taxon:4577"					
	/clone_lib="PCO124784"					
	/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"					
	[note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"]					
BASE COUNT	128 a	173 c	162 g	90 t		
ORIGIN						
Query Match	70.1%;	Score 474.2;	DB 11;	Length 553;		
Best Local Similarity	94.3%;	Pred. No. 8.5e-43;				
Matches	516;	Conservative	0;	Mismatches 23;	Indels 8;	Gaps 2;
QY	25	GCAGCAATCCACAAAGCCTTCGAAGGACCACACTG-----CTCGGAGGACACACCAAGCG	79			
Db	7	GCAGCAATCCACAAAGCCTTCGACGTTCGACGGGGGCTGCACACAGACACCAAGCG	66			
QY	80	TTGTGACCAATGGCTTACTACAGAGGTGACTCTCTCGGAGGAGTGAGTCTCGGTG	139			
Db	67	TGGGACCAATGGCTTACTACAGAGGTGACTCTCTCGGAGGAGTGAGTCTCGGTG	126			
QY	140	GCCTCGGCGGGCTTCGGCGGCCGACGGCGGGCTCCAGCAGCAGCTGCTCAAGGAGAAG	199			
Db	127	GCCTCGGCGGGCTTCGGCGGCCGACGGAGGGGGCTCCAGCAGCAGCTGCTCAAGGAGAAG	186			
QY	200	TTGAGAGGTTCGACAGGTATCACGGCGCGCGGCA--ACCACACACCATCATGGT	256			
Db	187	TTGAGAGGTTCGACAGGTCTCACGGCGCGCGGCAACACACCATCATGGT	246			
QY	257	CACACGGGGGCGCACGGCTTCGTGGTCGCGAGACCGAGGTCGAGGAGGACATCAACAC	316			
Db	247	CACACGGGGGCGCACGGCTTCGTGGTCGCGAGACCGAGGTCGAGGAGGACATCAACAC	306			
QY	317	TGACGGCGGAGGTCCACAGGCGCAGGAGAGCTTCCTCGCAGGCTTAAGTAGCGCGCC	376			
Db	307	TGACGGCGGAGGTCCACAGGCGCAGGAGAGCTTCCTCGCAGGCTTAAGTAGCGCGCC	366			
QY	377	CGGCGGCGGCATCCACGCGCGTTCGTGGTCGCGGCGCTTATGATGCTGTGTT	436			
Db	367	CGGCGGCGGCATCCACGCGCGTTCGTGGTCGCGGCGCTTATGATGCTGTGTT	426			
QY	437	GACTGTTGTGCGAGGTTCATCTGCTTGGCTTACGTGACGCGCAGCTCAGCTCCTGTA	496			
Db	427	GACTGTTGTGCGAGGTTCATCTGCTTGGCTTACGTGACGCGCAGCTCAGCTCCTGTA	486			
QY	497	CGAATTACGACAATAAGCTCGTACCTGTAATAAACTTCTCGTATATATATACCTACA	556			
Db	487	CGAATTACGACAATAAGCTCGTACCTGTAATAAACTTCTCGTATATATATACCTAAAAA	546			
QY	557	TCAAAAA	563			
Db	547	AAAAAAA	553			

RESULT 11	AW787732	945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.	500 bp	mRNA	linear	EST 16-MAY-2000
LOCUS	AW787732	GI:7844510				
DEFINITION	AW787732	Zea mays.				
ACCESSION	AW787732	EST.				
VERSION	AW787732.1	GI:7844510				
KEYWORDS	Zeas mays.					
SOURCE	ORGANISM					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1	(bases 1 to 500)				
AUTHORS	Walbot,V.					
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 945002 row: E column: 06. Location/Qualifiers 1..500 /organism="Zea mays" /cultivar="W23" /db_xref="taxon:4577" /clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)" /tissue.type="tassel, kernal, silk, husk, root, leaf" /dev_stage="fully-grown" /lab_host="DH10B" [note="Organ: tassel, kernal, silk, husk, root, leaf; Vector: pGAD10; Site_1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."]					
BASE COUNT	97 a	159 c	156 g	88 t		
ORIGIN						
Query Match	69.6%;	Score 470.6;	DB 10;	Length 500;		
Best Local Similarity	99.2%;	Pred. No. 2.2e-42;				
Matches	473;	Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	67	GACACACCAAGCGTCTGCACCAATGGCTTACTTACCAGGAGGTGGACTACTCTCGGAGGA	126			
Db	13	GACACACCAAGCGTTCGGCACCATTGGCTTACTTACCAGGAGGTGGACTACTCTCGGAGGA	72			
QY	127	GGTGAGGTTCGGTGGCCCCCGCGGCTTCGGCGCCGACGGCGCGCTCCAGCAGCAGT	186			
Db	73	GGTGAGGTTCGGTGGCCCCCGCGGCTTCGGCGCCGACGGCGCGCTCCAGCAGCAGT	132			
QY	187	CGTCAAGGAGAACTTCAGGAGGTTCACACGGTATCACGCCGCCGCCAACCCACCA				

```

/notes="Organ: tassell, kernal, silk, husk, root, leaf;
Vector: pGD10; Site_: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
a      165 c      92 t

```

BASE COUNT	135 a	165 c	160 g	92 t	Query Match	72.9%	Score 492.6	DB 10	Length 552	
ORIGIN					Best Local Similarity	97.3%	Pred. No. 8.9e-45			
					Matches 501	Conservative 0	Mismatches 14	Indels 0	Gaps 0	
Qy	67	GACACACCAAGCGTCTCCACCAATGGCTTACTACCAAGAGGTGAGCTACTGCTCGGAGGA	126							
Db	26	GACACACCAAGCGTGGCAGCAATGGCTTACTACCAAGAGGTGAGCTACTGCTCGGAGGA	85							
Qy	127	GGTGAGTCTGCTGGCCCGCGCGCTTCGGCCGACGCGGCGGCGTCCAGCAGCACGT	186							
Db	86	GGTGAGTCTGCTGGCCCGCGCGCTTCGGCCGACGCGGCGGCGTCCAGCAGCACGT	145							
Qy	187	CGTCAAGGAGAAGTTGAGAGAGGTGACACGGGTATCACGCGCGCGGCGCCAAACCACCA	246							
Db	146	CGTCAAGGAGAAGTTGAGAGAGGTCTATACGGTCTCACGCGCGCGGCGCCAAACCACCA	205							
Qy	247	CCACCATGTTGTCACACGCGGCGCCACGGCTTCGTGGTGGCCGAGACACAGGTGCAGAGGA	306							
Db	206	CCACCATGTTGTCACACGCGGCGCCACGGCTTCGTGGTGGCCGAGACACAGGTGCAGAGGA	265							
Qy	307	CATCAACACCTGCACCGGGGAGGTCCACGAGCGCAGGGAGAGCTTCTCGCCACAGGCTAA	366							
Db	266	CATCAACACCTGCACCGGGGAGTTCACGAGCGCAGGGAGAGCTTCTCGCCACAGGCTAA	325							
Qy	367	CTGAGCGCGCGCGCGGCATCCACGCGCGTTCGTGCTTGCCTCGCTGCTTATGTAT	426							
Db	326	CTGAGCGCGCGCGCGGCATCCACGCGCGTTCGTGCTTGCCTCGCTGCTTATGTAT	385							
Qy	427	GTCGTGGTTGACTGGTTGGCAGGGTCATCGTACTTGGCTATCGTACCTGTCACGCACTC	486							
Db	386	GTCGTGGTTGACTGGTTGGCAGGGTCATCGTACTTGGCTATCGTACCTGTCACGCACTC	445							
Qy	487	AGTCTCTGTACGAATTTACGACAAATAGCTCGTACCTGAATAAACTCTTCGTAAATCT	546							
Db	446	AGTCTCTGTACGAATTTACGACAAATAGCTCGTACCTGAATAAACTCTTCGTAAATCT	505							
Qy	547	AATACCTACATCAAAAAAAAAAAAAAAAAAAAAA	581							
Db	506	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	540							

RESULT 9	DEFINITION	AM288876	524 bp	mrna	linear	EST 16-JAN-2000
LOCUS		AW288876				
		70709E07.x4 707 - Mixed adult tissues from Walbot lab (SK)				zea
ACCESSION		AW288876				
VERSION		AW288876				
KEYWORDS		EST.				
SOURCE		ZEAL				
ORGANISM		zea mays.				
		zea mays				
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC				
REFERENCE		1 (bases 1 to 524)				
AUTHORS		Walbot, V.				
TITLE		Maize ESTs from various cDNA libraries sequenced at Stanford University				
JOURNAL		Unpublished (1999)				
COMMENT		Contact: Walbot V				
		Department of Biological Sciences				
		Stanford University				
		855 California Ave, Palo Alto, CA 94304, USA				
		Tel: 650 723 2227				

```

Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.
Location/Qualifiers
1. .524
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
107 a 170 c 157 g 89 t 1 others
BASE COUNT
ORIGIN

```

[illegible]

RESULT 10					
AY104409					
LOCUS	AY104409	553 bp	mRNA	linear	HTC 25-MAY-2002
DEFINITION	Zea mays PCol24784 mRNA sequence.				
ACCESSION	AY104409				
VERSION	AY104409.1 GI:21207487				
KEYWORDS	HTC.				
SOURCE	Zea mays.				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC				

[illegible]

REFERENCE AUTHORS TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 523) Walbot, V.
JOURNAL COMMENT	Maize ESTs from various cDNA libraries sequenced at Stanford University Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 707049 row: E column: 04.

```

Email: walbot@stanford.edu
Plate: 707049 row: E column: 04.

FEATURES
source
Location/Qualifiers
1..523
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
125 a 158 c 149 g 91 t
BASE COUNT

```

Vector: pGAD10; Site-1: EcoRI; cDNA library from tully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned.

BASE COUNT	125 a	158 c	149 g	91 t
ORIGIN				
Query Match	73.7%; Score 498.2; DB 10; Length 523;			
Best Local Similarity	98.4%; Pred. No. 2.3e-45;			
Matches 503; Conservative	0; Mismatches 8; Indels 0; Gaps 0;			


```

Db      540 ACTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 599
Qy      604 AAAAAAA 610
        |||||
Db      600 AAAAAAA 606

RESULT 5
BQ619383
LOCUS
DEFINITION
    RNOSEQ7B02_SK.ab1 Salt stressed Zea mays roots cDNA library EST 27-JUN-2002
    mays cDNA clone RNOSEQ7B02_SK.ab1 similar to No homology, mRNA
sequence.
ACCESSION
    BQ619383
VERSION
    BQ619383.1 GI:21621377
KEYWORDS
    EST.
SOURCE
    Zea mays.
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
    clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
    Wang, H. and Bohnert, H.J.
    Genomics of plant stress tolerance
    Unpublished (2002)
    Contact: Mark Fredricksen
    Department of Plant Biology
    University of Illinois
    1201 W. Gregory Dr., Urbana, IL 61801, USA
    Tel: 2172655473
    Email: bohnertlab@life.uiuc.edu.
FEATURES
    source
        1. 648
        /organism="Zea mays"
        /db_xref="taxon:4577"
        /clone="RNOSEQ7B02_SK.ab1"
        /clone_lib="Salt stressed Zea mays roots cDNA library"
        /tissue_type="Roots"
        /dev_stage="2 weeks old"
        /note="Vector: pBluescript SK+; Stressed 24 hours at 150
        mM NaCl"
BASE COUNT    172 a   193 c   183 g   100 t

Query Match    75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

Qy      67 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 126
Db      60 GACACACCAAGCGTCGGCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 119
Qy      127 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGCGCACGGCGGCGGCTCCAGCAGCACGT 186
Db      120 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGCGCACGGCGGCGGCTCCAGCAGCACGT 186
Qy      187 CGTCAAGGAGAGTTCGAGGAGGTTCGACCGGTATACACGGCGGCGGCGGCTCCAGCAGCACGT 243
Db      180 CGTCAAGGAGAGTTCGAGGAGGTTCGACCGGTATACACGGCGGCGGCGGCTCCAGCAGCACGT 239
Qy      244 CCACACCACTGCTGCACCGCGGAGGTCCACGAGCGGAGAGCTTCCTCGCCAGGCG 363
Db      300 GGACATCAACACTGACCGCGGAGGTCCACGAGCGGAGAGCTTCCTCGCCAGGCG 359
Qy      364 TAACTAGCGCGCGCGCGGCGGATCCACGCGGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Db      360 TAACTAGCGCGCGCGCGGCGGATCCACGCGGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Qy      424 TATGCTGTGTTGACTGTTGTGCGAGGTCATCGTACTTGGCTATCGTACGTGCGACGCA 483

```

```

Db      420 TATGCTGTGTTGACTGTTGTGTCAGGTCATCGTACTTGGCTATCGTACGTGCGACGCA 479
Qy      484 CTCAGCTCCTGTACGAATTAGCAATTAAGCTGCTGACCTGAATAAACTTCTTCGTAAT 543
Db      480 CTCAGCTCCTGTACGAATTAGCAATTAAGCTGCTGACCTGAATAAACTTCTTCGTAAT 539
Qy      544 ACTAATACCTACATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 603
Db      540 ACTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 599
Qy      604 AAAAAAA 610
        |||||
Db      600 AAAAAAA 606

RESULT 6
BQ619390
LOCUS
DEFINITION
    RNOSEQ7B09_SK.ab1 Salt stressed Zea mays roots cDNA library EST 27-JUN-2002
    mays cDNA clone RNOSEQ7B09_SK.ab1 similar to No homology, mRNA
sequence.
ACCESSION
    BQ619390
VERSION
    BQ619390.1 GI:21621384
KEYWORDS
    EST.
SOURCE
    Zea mays.
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
    clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
    Wang, H. and Bohnert, H.J.
    Genomics of plant stress tolerance
    Unpublished (2002)
    Contact: Mark Fredricksen
    Department of Plant Biology
    University of Illinois
    1201 W. Gregory Dr., Urbana, IL 61801, USA
    Tel: 2172655473
    Email: bohnertlab@life.uiuc.edu.
FEATURES
    source
        1. 648
        /organism="Zea mays"
        /db_xref="taxon:4577"
        /clone="RNOSEQ7B09_SK.ab1"
        /clone_lib="Salt stressed Zea mays roots cDNA library"
        /tissue_type="Roots"
        /dev_stage="2 weeks old"
        /note="Vector: pBluescript SK+; Stressed 24 hours at 150
        mM NaCl"
BASE COUNT    172 a   193 c   183 g   100 t

Query Match    75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

Qy      67 GACACACCAAGCGTCTGCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 126
Db      60 GACACACCAAGCGTCGGCACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGA 119
Qy      127 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGCGCACGGCGGCGGCTCCAGCAGCACGT 186
Db      120 GGTGAGGTGCGTGGCCCGCGGCTTCGGCCGCGCACGGCGGCGGCTCCAGCAGCACGT 186
Qy      187 CGTCAAGGAGAGTTCGAGGAGGTTCGACCGGTATACACGGCGGCGGCGGCTCCAGCAGCACGT 243
Db      180 CGTCAAGGAGAGTTCGAGGAGGTTCGACCGGTATACACGGCGGCGGCGGCTCCAGCAGCACGT 239
Qy      244 CCACACCACTGCTGCACCGCGGAGGTCCACGAGCGGAGAGCTTCCTCGCCAGGCG 363
Db      300 GGACATCAACACTGACCGCGGAGGTCCACGAGCGGAGAGCTTCCTCGCCAGGCG 359
Qy      364 TAACTAGCGCGCGCGGCGGATCCACGCGGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Db      360 TAACTAGCGCGCGCGGCGGATCCACGCGGCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Qy      424 TATGCTGTGTTGACTGTTGTGCGAGGTCATCGTACTTGGCTATCGTACGTGCGACGCA 483

```



```

/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150
mm NaCl"
172 a 193 c 183 g 100 t
BASE COUNT

```

Query Match	75.28;	Score 508.6;	DB 14;	Length 648;
Best Local Similarity	96.9%;	Prod. No. 1.5e-46;		
Matches 530;	Conservative 0;	Mismatches 14;	Indels 3;	Gaps 1;
QY	67	GACACACAAAGCGTCTCAGCAATGGCTTACTACACAGAGGTGGACTACTGCTCGGAGGA	126	
Db	60	GACACACAAAGCGTCGCACCAATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGA	119	
QY	127	GGTGAGGTGCGTGGCCCGCGCGCTTCGGCCCGCCACGGCGGGCGGTCCAGCAGCACGCT	186	
Db	120	GGTGAGGTGCGTGGCCCGCGCGCTTCGGCCCGCCACGGAGGGCGGTCCAGCAGCACGCT	179	
QY	187	CGTCAAGGAGAAGTTCGAGGAGGTCGACACGGGTATCACGGCCGGCGGCCA--ACCACCA	243	
Db	180	CGTCAAGGAGAAGTTCGAGGAGGTCGACACGGTCTCACGGCCGGCGGCCAACCACACCA	239	
QY	244	CCACCACCATGTCACACCGCGGCCACGGCTTCGTGGTGGCGGAGACAGGGTCTGAGGA	303	
Db	240	CCACCACCATGTCACACCGCGGCCACGGCTTCGTGGTGGCGGAGACAGGGTCTGAGGA	299	
QY	304	GGACATCAACACTGCACCGGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	363	
Db	300	GGACATCAACACTGCACCGGCGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCAGGGC	359	
QY	364	TAACTGAGCGCCCGCGCGGCATCCAGCCGCTTCGCTTGGCTGGGTGCCCTTATG	423	
Db	360	TAACTGAGCGCCCGCGCGGCATCCAGCCGCTTCGCTTGGCTGGGTGCCCTTATG	419	
QY	424	TATGCTGTGGTTGACTGTTGTCAGGGTCATCGTACTTGGCTATCGTACGTGTCACGGA	483	
Db	420	TATGCTGTGGTTGACTGTTGTCAGGGTCATCGTACTTGGCTATCGTACGTGTCACGGA	479	
QY	484	CTCAGCTCTGTACGAATTTACGACATAAGCTCGTCACCTGATATAACTTCTTCGTAT	543	
Db	480	CTCAGCTCTGTACGAATTTACGACATAAGCTCGTCACCTGATATAACTTCTTCGTAT	539	
QY	544	ACTAATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA	603	
Db	540	ACTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA	599	
QY	604	AAAAAAA 610		
Db	600	AAAAAAA 606		

[illegible]

```

University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
1. .648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="RNOSE06C06_SK_ab1"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/notes="Vector: pBluescript SK+; Stressed 24 hours at 150
mm NaCl"
172 a 193 c 183 g 100 t
BASE COUNT

```

Query Match	75.2%;	Score 508.6;	DB 14;	Length 648;
Best Local Similarity	96.9%;	Prod. No. 1.5e-46;		
Matches 530;	Conservative 0;	Mismatches 14;	Indels 3;	Gaps 1;
QY	67	GACACCAAGCGTCTGCACCAATGGCTTACTACCAAGAGGTGACACTGCTCGGAGGA	126	
DB	60	GACACCAAGCGTCGGACCAATGGCTTACTACCAAGAGGTGACACTGCTCGGAGGA	119	
QY	127	GGTAGGTCGGTGCGCCCGCGCGCTTCGGCCGCCACGGCGCGCGGTCCAGCAGCAGCT	186	
DB	120	GGTAGGTCGGTGCGCCCGCGCGCTTCGGCCGCCACGGAGCGCGGTCCAGCAGCAGCT	179	
QY	187	CGTCAAGGAGAAGTTCGAGGAGGTCGACAGGTATCACGGCGCCGCGCCCA---ACCACCA	243	
DB	180	CGTCAAGGAGAAGTTCGAGGAGGTCGACAGGTCTCACGGCGCGCGCCCAACCAACCA	239	
QY	244	CCACCACCATGGTACCAACCGCGCCACGGCTTCGTGGCGGAGACACAGGTCGAGGA	303	
DB	240	CCACCACCATGGTACCAACCGCGCCACGGCTTCGTGGCGGAGACACAGGTCGAGGA	299	
QY	304	GGACATCAACACCTGACCCGCGGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCGAGGCG	363	
DB	300	GGACATCAACACCTGACCCGCGGAGGTCCACGAGCGCAGGAGAGCTTCCTCGCGAGGCG	359	
QY	364	TAAGTAGCCGCGCGCGCGCGGCATCCACGCCGTCGTGCTTCGCTGGGTGCTTCGTATG	423	
DB	360	TAAGTAGCCGCGCGCGCGGCATCCACGCCGTCGTGCTTCGCTGGGTGCTTCGTATG	419	
QY	424	TATGTCGTGGTTGACTGGTTGTGCAGGGTCATCGTACTTGGCTATCGTACGTGCACGCA	483	
DB	420	TATGTCGTGGTTGACTGGTTGTGCAGGGTCATCGTACTTGGCTATCGTACGTGCACGCA	479	
QY	484	CTCAGTCTCTGTACGAATTACGACAATAAGCTCGTGACCTGAATAAAGCTTCCTCGTAAT	543	
DB	480	CTCAGTCTCTGTACGAATTACGACAATAAGCTCGTGACCTGAATAAAGCTTCCTCGTAAT	539	
QY	544	ACTTAATACCTACATCAAA	603	
DB	540	ACTTAATAAA	599	
QY	604	AAAAAAAA	610	
DB	600	AAAAAAAA	606	

RESULT	3
BQ619318	
LOCUS	
DEFINITION	BQ619318 RNSQ6D01_SK.abl Salt stressed Zea mays roots cDNA library Zea mays CDNA clone RNSQ6D01_SK.abl similar to No homology, mRNA sequence.
ACCSSION	BQ619318
VERSION	BQ619318.1 GI:21621312
KEYWORDS	EST.
SOURCE	Zea mays.
ORGANISM	Zea mays

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 01:23:01 ; Search time 1439.61 Seconds
(without alignments)
7604.943 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 accacgcgtcccccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_estl:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rtd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	508.6	75.2	648	14	BQ619167
2	508.6	75.2	648	14	BQ619315
3	508.6	75.2	648	14	BQ619318
4	508.6	75.2	648	14	BQ619337
5	508.6	75.2	648	14	BQ619383
6	508.6	75.2	648	14	BQ619390

7	498.2	73.7	523	10	AW331212
8	492.6	72.9	552	10	BE025302
9	481.2	71.2	524	10	AW288876
10	474.2	70.1	553	11	AY104409
11	470.6	69.6	500	10	AW787732
12	464.6	68.7	546	9	AA979839
13	462.8	68.5	550	12	BG840383
14	460.0	68.0	539	9	AI855425
15	456.4	67.5	496	10	BE129644
16	449.0	66.4	476	10	BE129644
17	449.0	66.4	481	9	AI964458
18	441.6	65.3	465	10	BE129897
19	429.6	63.6	436	10	AW787315
20	415.6	61.5	466	9	AI964534
21	409.0	60.5	443	10	BE129644
22	404.2	59.8	433	10	BE225008
23	361.2	53.4	401	10	AW289056
24	359.0	53.1	417	13	EM501439
25	323.6	47.9	541	10	AW745400
26	321.6	47.6	537	10	AW745400
27	319.6	47.3	326	10	BE025303
28	317.0	46.9	516	10	BE364814
29	316.0	46.7	519	10	AW680016
30	315.2	46.6	513	14	BQ280709
31	311.6	46.1	509	14	BQ280894
32	310.6	45.9	485	10	BE599123
33	310.6	45.9	522	10	AW679969
34	294.8	43.6	466	10	BE597738
35	275.4	40.7	362	12	BG840882
36	269.0	39.8	311	12	BF729420
37	247.2	36.3	399	13	BM318672
38	245.2	36.3	367	10	AW747603
39	226.8	33.6	348	10	BE593507
40	214.2	31.7	288	9	AI649804
41	212.6	31.4	274	10	AW288875
42	212.6	31.4	338	10	AW924579
43	185.4	27.4	512	14	BQ283209
44	185.4	27.4	583	14	BQ281558
45	183.6	27.2	277	10	BE593304

ALIGNMENTS

RESULT 1
BQ619167
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BQ619167
RNOSEQ4E05_SK.abi
mays cdna clone RNOSEQ4E05_SK.abi similar to No homology, mRNA
sequence.
BQ619167
RNOSEQ4E05_SK.abi
EST.
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
location/Qualifiers
l. . 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ4E05_SK.abi"

BQ619167
RNOSEQ4E05_SK.abi
mays cdna clone RNOSEQ4E05_SK.abi similar to No homology, mRNA
sequence.
BQ619167
RNOSEQ4E05_SK.abi
EST.
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
location/Qualifiers
l. . 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ4E05_SK.abi"

linear EST 27-JUN-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:52:33 ; Search time 29 Seconds
(without alignments)
331.081 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510
Sequence: 1 MAYQEVYDVCSEVRVAPV.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	510	100.0	93	9 US-10-090-035-2	Sequence 2, Appli
2	510	100.0	93	9 US-10-090-035-4	Sequence 4, Appli
3	510	100.0	93	9 US-10-090-035-6	Sequence 6, Appli
4	507	99.4	93	9 US-10-090-035-10	Sequence 10, Appli
5	499.5	97.9	94	9 US-10-090-035-8	Sequence 8, Appli
6	499.5	97.9	94	9 US-10-090-035-16	Sequence 16, Appli
7	295	57.8	91	9 US-10-090-035-18	Sequence 18, Appli
8	277.5	54.4	91	9 US-10-090-035-22	Sequence 22, Appli
9	273	53.5	92	9 US-10-090-035-20	Sequence 20, Appli
10	258	50.6	92	9 US-10-090-035-24	Sequence 24, Appli
11	245	48.0	99	9 US-10-090-035-14	Sequence 14, Appli
12	80	15.7	359	9 US-10-232-563-7	Sequence 7, Appli
13	80	15.7	359	9 US-09-888-615-64	Sequence 64, Appli
14	77	15.3	332	9 US-10-232-563-2	Sequence 2, Appli
15	77	15.1	359	9 US-10-232-563-6	Sequence 6, Appli
16	76	14.9	400	9 US-09-879-312-2	Sequence 2, Appli
17	73.5	14.4	285	9 US-10-278-173-16	Sequence 16, Appli
18	73.5	14.4	309	12 US-10-052-798-9	Sequence 9, Appli
19	73.5	14.4	312	12 US-10-052-798-10	Sequence 10, Appli

20	73	14.3	19	9 US-09-876-904A-627	Sequence 627, Appli
21	73	14.3	633	10 US-09-824-735-3	Sequence 3, Appli
22	73	14.3	633	10 US-09-801-368-338	Sequence 338, Appli
23	71.5	14.0	310	12 US-10-052-798-11	Sequence 11, Appli
24	71.5	14.0	354	10 US-09-925-302-501	Sequence 501, Appli
25	71.5	14.0	532	10 US-09-833-790-428	Sequence 428, Appli
26	69.5	13.6	315	9 US-10-232-563-11	Sequence 11, Appli
27	69.5	13.6	623	9 US-10-108-605-125	Sequence 125, Appli
28	69.5	13.6	623	9 US-10-108-605-129	Sequence 129, Appli
29	69.5	13.6	726	9 US-09-932-257A-19	Sequence 19, Appli
30	69	13.4	257	9 US-09-738-626-6135	Sequence 6135, Appli
31	68.5	13.4	120	9 US-09-051-013-3	Sequence 3, Appli
32	68	13.3	139	10 US-09-813-820-8	Sequence 8, Appli
33	67.5	13.2	118	10 US-09-205-658-120	Sequence 120, Appli
34	67.5	13.2	530	9 US-10-044-692-317	Sequence 317, Appli
35	67.5	13.2	530	9 US-10-044-533-317	Sequence 317, Appli
36	67.5	13.2	567	9 US-10-270-333-126	Sequence 126, Appli
37	67.5	13.2	605	9 US-09-741-233A-2	Sequence 2, Appli
38	67.5	13.2	1207	9 US-10-108-605-71	Sequence 71, Appli
39	67	13.1	446	10 US-09-853-386-69	Sequence 69, Appli
40	67	13.1	507	9 US-09-795-927-10	Sequence 10, Appli
41	67	13.1	507	9 US-09-738-897-2	Sequence 2, Appli
42	67	13.1	608	9 US-09-975-719-201	Sequence 201, Appli
43	67	13.1	639	9 US-09-975-719-200	Sequence 200, Appli
44	67	13.1	643	9 US-09-975-719-199	Sequence 199, Appli
45	67	13.1	643	10 US-09-853-386-70	Sequence 70, Appli

ALIGNMENTS

RESULT 1

US-10-090-035-2
; Sequence 2, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 357118/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2, 227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYQEVYDVCSEVRVAPVAGFGRGGVQHVYVKEFEVDTVSRAGANHHHHHHGG 60
DB 1 MAYQEVYDVCSEVRVAPVAGFGRGGVQHVYVKEFEVDTVSRAGANHHHHHHHHGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
DB 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 2

US-10-090-035-4
; Sequence 4, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 357118/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2, 227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

us-10-090-035-2.rapb

Mon Jun 16 14:55:54 2003

```

; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-4

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
Db 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 3
US-10-090-035-6
; Sequence 6, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-6

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
Db 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 4
US-10-090-035-10
; Sequence 10, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-10

Query Match
Best Local Similarity 99.4%; Score 507; DB 9; Length 93;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
Db 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
QY 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 5
US-10-090-035-8
; Sequence 8, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-8

Query Match
Best Local Similarity 97.9%; Score 499.5; DB 9; Length 94;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 59
Db 1 MAYIQEVDYCSSEVRSVAPAGFGRHGGVQGHVVKKEFEVDTVSRAGANHHHHHHHGG 60
QY 60 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 6
US-10-090-035-18
; Sequence 18, Application US/100900035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-18

```

; ORGANISM: Triticum aestivum
US-10-090-035-18

Query Match 97.9%; Score 499.5; DB 9; Length 94;
Best Local Similarity 98.9%; Pred. No. 5.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAYTQEVYDCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDVTSRAGAN-HHHHHHHG 59
Db 1 MAYTQEVYDCSEEVRSVAPAGFGRHGGVQOHHVVKKEFEEVDVTSRAGANHHHHHHHHG 60

Qy 60 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 7

US-10-090-035-16
; Sequence 16, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-090-035-16

Query Match 57.8%; Score 295; DB 9; Length 91;
Best Local Similarity 60.6%; Pred. No. 4e-23;
Matches 57; Conservative 14; Mismatches 19; Indels 4; Gaps 2;

Qy 1 MAYTQEVYDCSEEVRSVAP-AGFGRHGGVQOHHVVKKEFEEVDVTSRAGANHHHHHHHHG 59
Db 1 MAHQEVYDCSEEVRSVTPGTGGFLGRGGVQOHHVVKKEFEEVDVTSRAGANHHHHHHHHG 57

Qy 60 GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 58 NDYLMVRTRVEEDINTCTGEFRKQSFLLKSD 91

RESULT 8

US-10-090-035-22
; Sequence 22, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-22

Query Match 54.4%; Score 277.5; DB 9; Length 91;
Best Local Similarity 61.2%; Pred. No. 2.4e-21;

Matches 60; Conservative 12; Mismatches 13; Indels 13; Gaps 6;
Qy 1 MAYTQEVYDCSEEVRSVAPAGFGRHGGVQOHHVVKKEF-EEVDVTSRAGANHHHHHHHHG 58
Db 1 MAHQEVYDCSEEVRSVAGNPA---RRGGVQOHHVVKKEFVQEFDTAGR----RHGHHGH 53
Qy 59 G---GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 54 GRGSGH-FFEVRESLEEDINTRTGEFHERKENFVVRAD 90

RESULT 9

US-10-090-035-20
; Sequence 20, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-20

Query Match 53.5%; Score 273; DB 9; Length 92;
Best Local Similarity 61.6%; Pred. No. 7e-21;
Matches 61; Conservative 12; Mismatches 12; Indels 14; Gaps 7;

Qy 1 MAYTQEVYDCSEEVRSVAPAGFGRHGGVQOHHVVKKEF-EEVDVTSRAGANHHHHHHHHG 57
Db 1 MAHQEVYDCSEEVRSVAGVPA---RRCGGVQOHHVVKKEFVQEFDTAGR----RHGHHGH 53

Qy 58 HG---GHGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 54 HGRGSGH-FFEVRESLEEDINTRTGEFHERKENFVVRAD 91

RESULT 10

US-10-090-035-24
; Sequence 24, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-24

Query Match 50.6%; Score 258; DB 9; Length 92;
Best Local Similarity 58.6%; Pred. No. 2.4e-19;
Matches 58; Conservative 14; Mismatches 13; Indels 14; Gaps 7;

Qy 1 MAYTQEVYDCSEEVRSVAPAGFGRHGGVQOHHVVKKEF-EEVDVTSRAGANHHHHHHHHG 57
Db 1 MAHQEVYDCSEEVRSVAGVPA---RRCGGVQOHHVVKKEFVQEFDTAGR----RHGHHGH 53

!


```

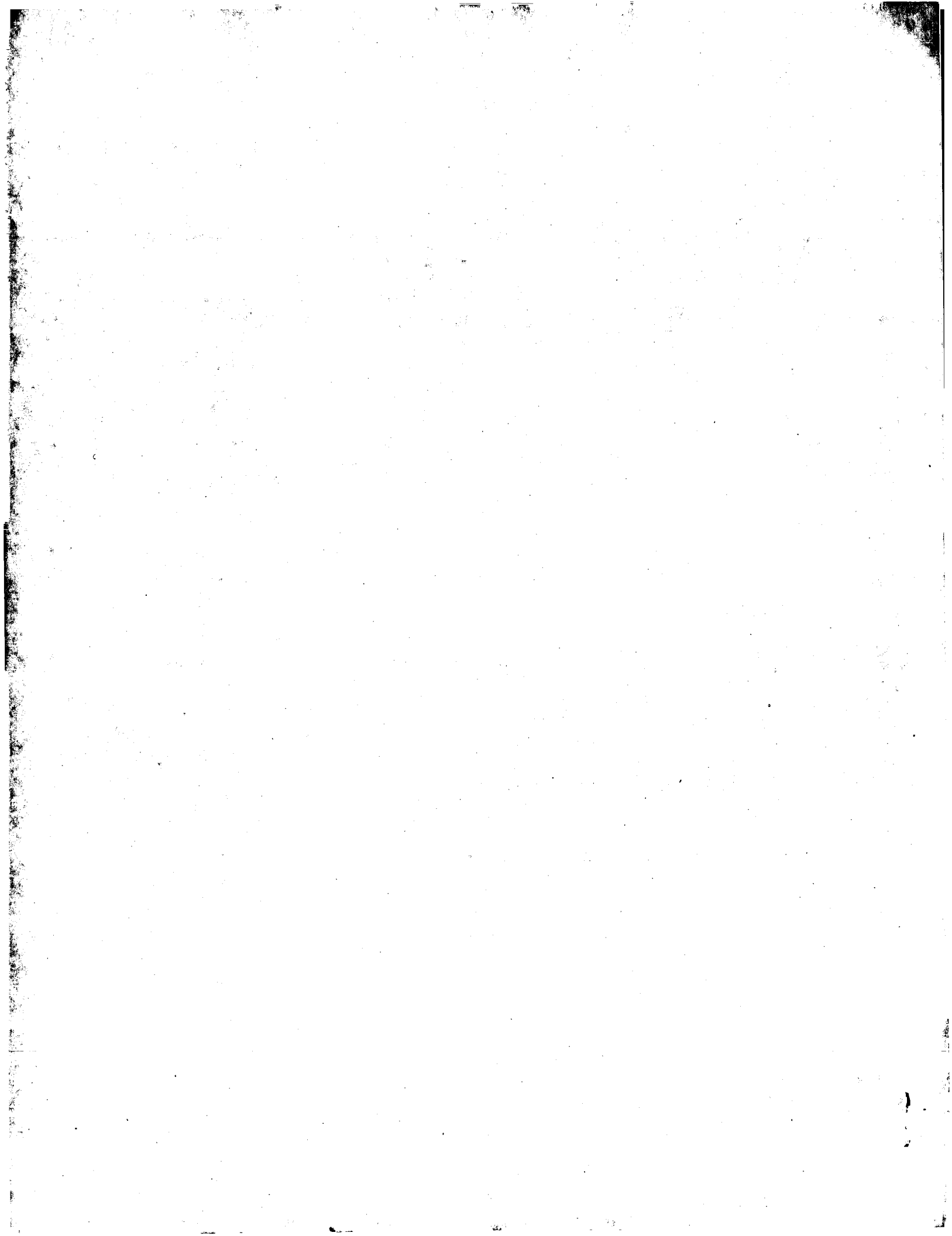
Db      230 QLVSMVRELNQLRGF 246
      :: |:: |
RESULT 15
US-10-232-563-6
; Sequence 6, Application US/10232563
; Publication No. US20030087394A1
; GENERAL INFORMATION:
; APPLICANT: Sharma, Arun
; TITLE OF INVENTION: INSULIN RELATED TRANSCRIPTION FACTOR AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 10276-072001
; CURRENT APPLICATION NUMBER: US/10/232,563
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/316,453
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-232-563-6

Query Match      15.1% Score 77; DB 9; Length 359;
Best Local Similarity 25.0%; Pred. No. 2.8;
Matches 24; Conservative 7; Mismatches 29; Indels 36; Gaps 4;

QY      1 MAYQOEVGYCSEEVRSVAP-----AGFGRHGG--GVQQHVVKKEFEVDTVSRAGA-- 49
Db      117 MSGYQH--HLNPEALNLTPEDAVEALIGSGHHGAHHGHPAAAAAYEAFRGQSFAGGG 174
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      50 -----NHHHHHHHHGGHG 62
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      175 ADDMGAGHHGCAHTAHHHSAHHHHHHHHHHHHHHHHHHHHHHHHHHHHGGSG 210

```

Search completed: June 6, 2003, 13:02:10
Job time : 30 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:20:30 ; Search time 39 Seconds
(without alignments)
731.306 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYQVEDYCEEVRSVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US10090035/runat_06062003_105504_10985/app_query.fasta_1.526
-DB=issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEA SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10090035 -CCGN=1.1.55 -runat_06062003_105504_10985 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA :
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq : *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq : *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq : *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq : *
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq : *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	83	16.3	4403765	4	US-09-103-840A-2
C 2	83	16.3	4411529	4	US-09-103-840A-1
C 3	80.5	15.8	2353	5	PCT-US92-06840-1
C 4	80	15.7	2310	3	US-09-036-987A-25
C 5	80	15.7	2310	4	US-09-370-700-25
C 6	79.5	15.6	195	2	US-08-622-740-9
C 7	79.5	15.6	195	3	US-08-440-689-9
C 8	79.5	15.6	195	4	US-09-122-399-9
C 9	77	15.1	2373	4	US-08-789-275-1
C 10	77	15.1	3218	2	US-08-677-862-1
C 11	77	15.1	3218	2	US-09-252-571-1
C 12	77	15.1	3218	3	US-09-434-065-1

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.22e+04
Score: 83.00
Percent Similarity: 46.55%
Best Local Similarity: 36.21%
Query Match: 16.27%
DB: 4
Length: 4403765
Matches: 21
Conservative: 6
Mismatch: 29
Indels: 2
Gaps: 1

US-10-090-035-2 (1-93) x US-09-103-840A-2 (1-4403765)

[illegible]

Db 483 GCCCATGATCGCTCTGCAGCGCTGTCACCGAGGAC-----CCGACCGAGGT 530
 QY 81 LHisGluArgArgGluSerPheLeuAlaArg 91
 Db 531 GCACCAACCAGGAGGTGATCTCGTGCAG 561

RESULT 4

US-09-036-987A-25/c
 ; Sequence 25, Application US/09036987A
 ; Patent No. 6143526
 ; GENERAL INFORMATION:
 ; APPLICANT: Baltz, Richard H.
 ; APPLICANT: Broughton, Mary C.
 ; APPLICANT: Crawford, Kathryn P.
 ; APPLICANT: Madduri, Krishnamurthy
 ; APPLICANT: Merlo, Donald J.
 ; APPLICANT: Treadway, Patti J.
 ; APPLICANT: Turner, Jan R.
 ; APPLICANT: Waldron, Clive
 ; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 ; TITLE OF INVENTION: Production
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dow Agrosciences LLC Patent Department
 ; STREET: 9330 Zionsville Road
 ; CITY: Indianapolis
 ; STATE: Indiana
 ; COUNTRY: USA
 ; ZIP: 46268

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/036.987A
 FILING DATE: 09-MAR-1998
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R
 REGISTRATION NUMBER: 28,479
 REFERENCE/DOCKET NUMBER: 50,608
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (317)337-4816
 TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2310 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 88..1077
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1165..1992

US-09-036-987A-25

Alignment Scores:

Pred. No.: 1.24 Length: 2310
 Score: 80.00 Matches: 27
 Percent Similarity: 41.76% Conservative: 11
 Best Local Similarity: 29.67% Mismatches: 35
 Query Match: 15.69% Indels: 18
 DB: 3 Gaps: 4

US-10-090-035-2 (1-93) x US-09-036-987A-25 (1-2310)

QY 7 ValAspTyrCysSerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGly 26
 Db 480 GTAGACCTCGTGGTGGAAACGTGCACGAACCTTGCAGTCTCGCGCTCGAGCGC----- 427

QY 27 GlyGlyValGlnGlnHisValValLysGluLysPheGluValValAspThrValSerArg 46
 Db 426 ---GGCCTGCGACGACGACGTTGGTGGCGGACACCGTGGTGTATCACGAAGCGTC---CGA 373
 QY 47 AlaGly-----AlaAsnHisHisHisHisHisHis 55
 Db 372 GCCGGTGTATCGAGCGGTGCGAGTGGTTCGGCGGCGAAGTGCACCAACGCTCCACGCC 313
 QY 56 GlyHisHisGlyGlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsn 75
 Db 312 GGACATCAGCGCGCAACACGATTCGCGTTCGAGATGTCGCCCGACGACGAA----- 262
 QY 76 ThrCysThrGlyGluValHisGluArgGlu 86
 Db 261 CTTACGCGGGGGTGTTCGCGGACGCGCGCCAG 229

RESULT 5

US-09-370-700-25/c
 ; Sequence 25, Application US/09370700
 ; Patent No. 6274350
 ; GENERAL INFORMATION:
 ; APPLICANT: Baltz, Richard H.
 ; APPLICANT: Broughton, Mary C.
 ; APPLICANT: Crawford, Kathryn P.
 ; APPLICANT: Madduri, Krishnamurthy
 ; APPLICANT: Treadway, Patti J.
 ; APPLICANT: Turner, Jan R.
 ; APPLICANT: Waldron, Clive
 ; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 ; FILE REFERENCE: 50489 DIV1
 ; CURRENT APPLICATION NUMBER: US/09/370,700
 ; CURRENT FILING DATE: 1999-08-09
 ; EARLIER APPLICATION NUMBER: US 09/36987
 ; EARLIER FILING DATE: 1998-03-09
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 25
 ; LENGTH: 2310
 ; TYPE: DNA
 ; ORGANISM: Saccharopolyspora spinosa
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (88)..(1077)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1165)..(1992)
 ; US-09-370-700-25

Alignment Scores:

Pred. No.: 1.24 Length: 2310
 Score: 80.00 Matches: 27
 Percent Similarity: 41.76% Conservative: 11
 Best Local Similarity: 29.67% Mismatches: 35
 Query Match: 15.69% Indels: 18
 DB: 4 Gaps: 4

US-10-090-035-2 (1-93) x US-09-370-700-25 (1-2310)

QY 7 ValAspTyrCysSerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGly 26
 Db 480 GTAGACCTCGTGGTGGAAACGTGCACGAACCTTGCAGTCTCGCGCTCGAGCGC----- 427
 QY 27 GlyGlyValGlnGlnHisValValLysGluLysPheGluValValAspThrValSerArg 46
 Db 426 ---GGCCTGCGACGACGACGTTGGTGGCGGACACCGTGGTGTATCACGAAGCGTC---CGA 373
 QY 47 AlaGly-----AlaAsnHisHisHisHisHisHis 55
 Db 372 GCCGGTGTATCGAGCGGTGCGAGTGGTTCGGCGGCGAAGTGCACCAACGCTCCACGCC 313
 QY 56 GlyHisHisGlyGlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsn 75

Sequence 9, Application US/08440689
Patent No. 6025545
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
Production of Stably Transformed, Fertile Monocot Plants
TITLE OF INVENTION: and Cells Thereof
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/440,689
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.013US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-440-689-9

Alignment Scores:
Pred. No.: 0.0531 Length: 195
Score: 79.50 Matches: 23
Percent Similarity: 48.98% Conservative: 1
Best Local Similarity: 46.94% Mismatches: 13
Query Match: 15.59% Indels: 12
DB: 3 Gaps: 3

US-10-090-035-2 (1-93) x US-08-440-689-9 (1-195)

QY 16 SerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValVallys 35
Db 153 TCTGGAGGACCGCGGGCGGCGGAGGCTGGCGTGGACTT-----112
QY 36 GluLysPheGluValAspThrValSerArgAlaGlyAlaAsnHisHisHis 55
Db 111 GAGCCCTCGAAGCGGACGCGGCTGGCGGACGAGGC-----CATCATCAC 64

QY 56 Gly-HisHisGlyGly---HisGly 62
Db 63 GGTTCATCATCGGTGGCGCCATGGC 39

RESULT 8
US-09-122-399-9/c
Sequence 9, Application US/09122399
Patent No. 6329574
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
Production of Stably Transformed, Fertile Monocot Plants

312 GGACATCAGCGCGCCCAACAGTTCGCGTTCGAGATGTCGCGCGACAGCAA-----262
QY 76 ThrCysThrGlyGluValHisGluArgArgGlu 86
Db 261 CTTACGCGGGGTGTTCGCGACCGCGCCAG 229

RESULT 6
US-08-622-740-9/c
Sequence 9, Application US/08622740
Patent No. 590390
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
Production of Stably Transformed, Fertile Monocot Plants
TITLE OF INVENTION: and Cells Thereof
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/622,740
FILING DATE: 27-MARCH-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.013US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-622-740-9

Alignment Scores:
Pred. No.: 0.0531 Length: 195
Score: 79.50 Matches: 23
Percent Similarity: 48.98% Conservative: 1
Best Local Similarity: 46.94% Mismatches: 13
Query Match: 15.59% Indels: 12
DB: 3 Gaps: 3

US-10-090-035-2 (1-93) x US-08-622-740-9 (1-195)

QY 16 SerValAlaProAlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValVallys 35
Db 153 TCTGGAGGACCGCGGGCGGCGGAGGCTGGCGTGGACTT-----112
QY 36 GluLysPheGluValAspThrValSerArgAlaGlyAlaAsnHisHisHis 55
Db 111 GAGCCCTCGAAGCGGACGCGGCTGGCGGACGAGGC-----CATCATCAC 64

QY 56 Gly-HisHisGlyGly---HisGly 62
Db 63 GGTTCATCATCGGTGGCGCCATGGC 39

RESULT 7
US-08-440-689-9/c

PERMMENT REFLECTION NUMBER: US/08/189,275P

PERMANENT REGISTRATION NUMBER: US/08//89,275A

STRANDEDNESS: double
TOPOLOGY: linear

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

US-10-090-035-2 (1-93) x US-09-252-571-1 (1-3218)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
DB 1783 AGTCCCAGGTGCGTCAGCAATTCCTGCTCTGTTGGTTCAGGCACTGAAGCTCCT 1842
QY 28 -----GlyValGlnGlnHisValValLysGluLysPheGluGluValAspThrVal 44
DB 1843 ACACAGGTCACTGTTGAACCTCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisGly 56
DB 1903 CAGAATGCATGTCATCATCACCATGTTAACAAGTTCCTCATCACCACCAACCCAC 1962
QY 57 HisHisGlyGlyHisGly 62
DB 1963 CATCACCACCACCATGGA 1980

RESULT 12
US-09-434-065-1
; Sequence 1, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/434,065
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/677,862
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-434-065-1

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

Alignment Scores: 4.58 Length: 3218
Pred. No.: 77.00 Matches: 22
Score: 39.39% Conservative: 4
Percent Similarity: 33.33% Mismatches: 26
Best Local Similarity: 15.10% Indels: 14
Query Match: 2 Gaps: 2
DB: 2

US-10-090-035-2 (1-93) x US-08-677-862-1 (1-3218)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
DB 1783 AGTCCCAGGTGCGTCAGCAATTCCTGCTCTGTTGGTTCAGGCACTGAAGCTCCT 1842
QY 28 -----GlyValGlnGlnHisValValLysGluLysPheGluGluValAspThrVal 44
DB 1843 ACACAGGTCACTGTTGAACCTCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisGly 56
DB 1903 CAGAATGCATGTCATCATCACCATGTTAACAAGTTCCTCATCACCACCAACCCAC 1962
QY 57 HisHisGlyGlyHisGly 62
DB 1963 CATCACCACCACCATGGA 1980

RESULT 11
US-09-252-571-1
; Sequence 1, Application US/09252571
; Patent No. 5981250
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/252,571
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-252-571-1

Percent Similarity: 39.39% Conservative: 4
 Best Local Similarity: 33.33% Mismatches: 26
 Query Match: 15.10% Indels: 14
 DB: 3 Gaps: 2

US-10-090-035-2 (1-93) x US-09-434-065-1 (1-3218)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGly----- 27
 DB 1783 AGTCCCGAGGTGGTCACCAATTCCTCTTGTGGTCAGGCACTGAAGCTCCT 1842
 QY 28 -----GlyValGlnGlnHisValValLysGluLysPheGluGluValAspThrVal 44
 DB 1843 ACACAGGTCACTGTTGAACATCATCTGTTCAAGAAACAACCTTTCATGTAGCCCTCAA 1902
 QY 45 SerArgAla-----GlyAlaAsnHisHisHisHisHisHisGly 56
 DB 1903 CAGATGATGTCATCATCACCAGTGAACAGTTCCCATCACCACCACTCACCACCA 1962
 QY 57 HisHisGlyGlyHisGly 62
 DB 1963 CATCACCACCACTGGA 1980

RESULT 13

US-09-086-010-1
 ; Sequence 1, Application US/09086010
 ; Patent No. 6274338
 ; GENERAL INFORMATION:
 ; APPLICANT: GlmCher, Laurie H. et al.
 ; TITLE OF INVENTION: Human c-Maf Compositions and
 ; METHODS OF USE THEREOF
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA: HUI-027CP
 ; APPLICATION NUMBER: US/09/086,010
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/030,579
 ; FILING DATE: 2-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kara, Catherine J.
 ; REGISTRATION NUMBER: 41,106
 ; REFERENCE/DOCKET NUMBER: HUI-027CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)42-4214
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1203 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1203
 ; US-09-086-010-1

Alignment Scores:
 Pred. No.: 1.65 Length: 1203
 Score: 76.00 Matches: 18

Percent Similarity: 45.61% Conservative: 8
 Best Local Similarity: 31.58% Mismatches: 25
 Query Match: 14.90% Indels: 6
 DB: 4 Gaps: 1

US-10-090-035-2 (1-93) x US-09-086-010-1 (1-1203)

QY 11 SerGluGluValArgSerValAlaProAlaGlyPheGlyArgHisGlyGlyValGln 30
 DB 403 GCGCACAGCTAGCCGCGCGCGCGCGCGCTCTTGGCGGCGCAGCGGC 462
 QY 31 GlnHis-----ValValLysGluLysPheGluGluValAspThrVal 44
 DB 463 GAGGAGATGGCGCCGCCCGCGGTGTCGCGGTATCGCGCGCGCGCGCGCAG 522
 QY 45 SerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyHisGlyHis 61
 DB 523 ACGGCAC 573

RESULT 14

US-09-329-234A-6
 ; Sequence 6, Application US/09329234A
 ; Patent No. 6331416
 ; GENERAL INFORMATION:
 ; APPLICANT: ShanL, ziv
 ; APPLICANT: Shoseyov, Oded
 ; TITLE OF INVENTION: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND R
 ; TITLE OF INVENTION: PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTUR
 ; FILE REFERENCE: 00/20274
 ; CURRENT APPLICATION NUMBER: US/09/329,234A
 ; CURRENT FILING DATE: 2001-09-04
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 1305
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Construct containing Protein L, fused to CBDcex sequence
 ; US-09-329-234A-6

Alignment Scores:

Pred. No.: 3.27 Length: 1305
 Score: 74.00 Matches: 16
 Percent Similarity: 40.91% Conservative: 2
 Best Local Similarity: 36.36% Mismatches: 22
 Query Match: 14.51% Indels: 4
 DB: 4 Gaps: 1

US-10-090-035-2 (1-93) x US-09-329-234A-6 (1-1305)

QY 19 ProAlaGly-----PheGlyArgHisGlyGlyValGlnHisValVal 34
 DB 1173 CCG 1232
 QY 35 LysGluLysPheGluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHis 54
 DB 1233 CCGACGCGGTTCGCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1292
 QY 55 HisGlyHisHis 58
 DB 1293 CACCAACCAAC 1304

RESULT 15

US-08-240-783B-1/c
 ; Sequence 1, Application US/08240783B
 ; Patent No. 5756348
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Kelli
 ; APPLICANT: Borden, Laurence A.
 ; APPLICANT: Branchek, Theresa
 ; APPLICANT: Hartig, Paul R.
 ; APPLICANT: Weinshank, Richard L.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 11:41:43 ; Search time 125 Seconds
(without alignments)
1077.357 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYQYEDVCSEVRVAPV.....INTCTGEVHERESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xxlh
-MODEL-frame-p2n.model -USPTO.spool/US10090035/runat_06062003_105507_11116/app-query.fasta_1.526
-DB-publishedApplications NA -OFMT-fastap -SUFFIX-rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -SPARP-1 -END-1 -MATRIX-blosum62
-TRANS-human40.cdi -LIST-45 -DOALIGN=200 -THR_SCORE-pct -THR_MAX=100
-THR_MIN=0 -ALIGN-15 -MODS-LOCAL -OUTFMT-ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER-US10090035@cgn_1_176@runat_06062003_105507_11116
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	574	9	US-10-090-035-3
2	510	100.0	577	9	US-10-090-035-5
3	510	100.0	676	9	US-10-090-035-1
4	507	99.4	529	9	US-10-090-035-9
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 9, Appli

5	499.5	97.9	524	9	US-10-090-035-17
6	499.5	97.9	580	9	US-10-090-035-7
7	295	57.8	591	9	US-10-090-035-15
8	277.5	54.4	436	9	US-10-090-035-21
9	258	50.6	584	9	US-10-090-035-19
10	258	50.6	584	9	US-10-090-035-23
11	251	49.2	348	9	US-10-090-035-13
12	213	41.8	237	10	US-09-923-876-2788
13	92.5	18.1	328	10	US-09-924-035A-17
14	92.5	18.1	591	9	US-09-938-842A-1858
15	88.5	17.4	546	10	US-09-728-445-817
16	83	16.3	3870	9	US-09-712-363-113
17	82.5	16.2	3966	9	US-10-198-846-13254
18	80	15.7	10062	10	US-09-888-615-5
19	78.5	15.4	1602	10	US-09-833-790-417
20	78	15.3	1059	9	US-10-232-563-4
21	78	15.3	1189	10	US-09-900-237-5
22	78	15.3	2681	9	US-10-232-563-1
23	77	15.1	1080	9	US-10-232-563-5
24	77	15.1	6381	10	US-09-969-347-216
25	76	14.9	1203	9	US-09-879-312-1
26	76	14.9	2627	9	US-10-037-270-434
27	75.5	14.8	491	9	US-09-854-133-343
28	75.5	14.8	491	10	US-09-738-973-343
29	75.5	14.8	25309	9	US-10-291-737-3
30	75	14.7	8979	9	US-09-738-626-2739
31	75	14.7	3309400	9	US-09-738-626-1
32	74.5	14.6	1458	9	US-10-103-313-83
33	74.5	14.6	1519	9	US-10-153-668-95
34	74	14.5	464	10	US-09-833-790-30
35	74	14.5	558	9	US-09-998-660-1
36	74	14.5	593	10	US-09-864-761-6743
37	74	14.5	910	12	US-10-062-254-325
38	73.5	14.4	32082	9	US-09-764-891-9679
39	73.5	14.4	279	9	US-09-984-130-143
40	73.5	14.4	858	9	US-09-938-842A-1494
41	73.5	14.4	930	12	US-10-052-798-6
42	73.5	14.4	939	12	US-10-052-798-7
43	73.5	14.4	1084	9	US-10-278-173-15
44	73.5	14.4	2331	9	US-09-938-842A-457
45	73	14.3	472	9	US-09-918-995-6480

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Alignment Scores: 1.3e-58
Pred. No.: 510.00
Score: 93
Percent Similarity: 100.00%
Length: 574
Matches: 93
Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-3 (1-574)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 96 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 155
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 156 GGCTTCGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCGTCAAGAGAAGTTCAGGAG 215
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 216 GTGCACAGGTCTCACGCGCGCGCGGCCACACACCACCATGGTCCACACGCGCGC 275
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 276 CACGGCTTCGTGGCGCGAGACGAGGTGCGAAGAGGACATCAACACCTGCACCGCGG 335
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 336 GTCCACGAGCGCAGGAGAGCTTCTTCGCCAGGGCTAAC 374

RESULT 2
US-10-090-035-5
; Sequence 5, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(377)
US-10-090-035-5
Alignment Scores:
Pred. No.: 1.31e-58 Length: 577
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-5 (1-577)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 99 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 158
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 159 GGCTTCGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCGTCAAGAGAAGTTCAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 219 GTGCACAGGTCTCACGCGCGCGCGGCCACACACCACCATGGTCCACACGCGCGC 278
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80

RESULT 3
US-10-090-035-1
; Sequence 1, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(367)
US-10-090-035-1
Alignment Scores:
Pred. No.: 1.59e-58 Length: 676
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-2 (1-93) x US-10-090-035-1 (1-676)

QY 1 MetAlaTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 89 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 148
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 149 GGCTTCGGCGCCACGCGCGCGCTCCAGCAGCAGTCTCGTCAAGAGAAGTTCAGGAG 208
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 209 GTGCACAGGTATCACGCGCGCGCGCCACACACCACCATGGTCCACACGCGCGC 268
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 269 CACGGCTTCGTGGTGGCGGAGACGAGGTGCGGAGGAGGACATCAACACCTGCACCGCGG 328
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 329 GTCCACGAGCGCAGGAGAGCTTCTTCGCCAGGGCTAAC 367

RESULT 4
US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 60/272,227
```

```

; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 529
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(311)
US-10-090-035-q

```

Alignment Scores:	
Pred. No.:	2,93e-58
Score:	507.00
Percent Similarity:	100.00%
Best Local Similarity:	98.92%
Query Match:	99.41%
DB:	9
Length:	539
Matches:	92
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-090-035-2 (1-93) x US-10-090-035-9 (1-529)

QY	1	MetAlaTyrTyrGlnGluValuAspTyrCysSerGluGluValArgSerValAlaProAla	20
Db	53	ATGGCTTACTACACGAGGTGGACTACTGTCGAGGAGTGAGTGGTGGCCCGGCC	112
QY	21	GlyPheGlyArgHisGlyGlyGlyValClnGlnHisValValLysGluLysPheGluGlu	40
Db	113	GGCTTCGGCCGCCACCGCGCGCTCCACGACGCTGTCGAAGGAGAGTTCGAGGAG	172
QY	41	ValAspThrValSerArgAlaGlyAlaAsnHisHisHisGlyHisGlyGlyGly	60
Db	173	GTCCAGACGGTTCGACGCGCGCGGCCCAACACACACCATGGTCACCACGCGGC	232
QY	61	HisGlyPheValValArgGluThrArgValClnGluAspIleAsnThrCysThrGlyGlu	80
Db	233	CACGGCTTCGTGGTGGCGGAGACCGAGGTTCGAGGAGCATCAACACCTGCACCGCGGAG	292
QY	81	ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn	93
Db	293	GTCCACGAGCGCAGGAGAGCTTCCTCGCCAGCGCTAAC	331

```

RESULT 5
US-10-090-035-17
: Sequence 17, Application US/10090035
: Patent No. US20020170089A1
: GENERAL INFORMATION:
: APPLICANT: Simmons, Carl R.
: TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
: TITLE OF INVENTION: Proteins and Uses Thereof
: FILE REFERENCE: 35718/242990
: CURRENT APPLICATION NUMBER: US/10/090,035
: CURRENT FILING DATE: 2002-02-28
: PRIOR APPLICATION NUMBER: 60/272,227
: PRIOR FILING DATE: 02/28/2001
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 524
: TYPE: DNA
: ORGANISM: Triticum aestivum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (57)...(338)
: NAME/KEY: misc_feature
: LOCATION: (1)...(524)
: OTHER INFORMATION: n = A,T,C or G
JS-10-090-035-17

```

Alignment Scores:	2.85e-57	524
red. NO.:	499.50	93
Score:	98.94%	0
Percent Similarity:		
Length:		
Matches:		
Conservative:		

Best Local Similarity:	98.94%	Mismatches:	0
Query Match:	97.94%	Indels:	1
DB:	9	Gaps:	1

US-10-090-035-2 (1-93) x US-10-090-035-17 (1-524)

QY	1	MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla	20
Db	57	ATGGCTTACTACAGAGGTGGACTACTGCTCGGAGGAGTGAGGTGGTGGCCGGCC	116
QY	21	GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu	40
Db	117	GGCTTCGGCGCCACGGAGGGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAG	176
QY	41	ValAspThrValSerArgAlaGlyAlaAsn--HisHisHisHisGlyHisHisGly	59
Db	177	GTCGACACGGTCTCAGCGCGGGCGCCAAACCACCACCACCACCATGTGTCAACACGCG	236
QY	60	GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly	79
Db	237	GGCCACGGCTTCGTGGTGGCGAGACCACGAGGTTCGAGGAGACATCAACACTGCACCGCG	296
QY	80	GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn	93
Db	297	GAGGTTCACGAGCGCAGGAGAGGTTCCTCGCCAGGGCTAAC	338

RESULT 6

```

US-10-090-035-7
; Sequence 7, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(380)
US-10-090-035-7

```

Alignment Scores:		
Pred. No.:	3.24e-57	580
Score:	499.50	Length:
Percent Similarity:	98.94%	Matches:
Best Local Similarity:	98.94%	Conservative:
Query Match:	97.94%	Mismatches:
DB:		Indels:
		Gaps:

US-10-090-035:2 (1-93) x US-10-090-035-7 (1-580)

[illegible]

```
Db      279  GGCCACGGCTTCGTGTCGGCAGACAGGCTCGAGGAGGACATCAACACCTGCACCGGC 338
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      339  GAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGGCTAAC 380

RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.:      4e-30      Length:      591
Score:          295.00     Matches:      57
Percent Similarity: 75.53%  Conservative: 14
Best Local Similarity: 60.64% Mismatches:    19
Query Match:     57.84%   Indels:        4
DB:              9       Gaps:          2
US-10-090-035-2 (1-93) x US-10-090-035-15 (1-591)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db      61  ATGGCTACTACCAAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGACCCACC 120
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db      121  GGCGGCTTCCTCGCGCGCGCGGCGTGCAGCAGCAGCAGCTCGTCAAGAGAGAGCTTCCAG 180
QY      40  GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db      181  GAGATCGAC-----AGTCCGGCTCCGGCGCCGCCACCAACCAACCAACCAACCAAC 231
QY      60  GlyHisGlyPheValArgGluThrArgValGluGluAspPheLeuAlaArgAlaAsn 79
Db      232  AACGACTACCTGATGGTGGCGGAGCAAGAGTGGAGGAGGAGTCAACACCTGCACCGGC 291
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      292  GAGTTCGCGAGCGGAGCAGAGCTTCCTGCTCAAGTCCGAC 333

RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:
Pred. No.:      5.68e-28      Length:      436
Score:          277.50     Matches:      60
Percent Similarity: 73.47%  Conservative: 12
Best Local Similarity: 61.22% Mismatches:    13
Query Match:     54.41%   Indels:        13
DB:              9       Gaps:          6
US-10-090-035-2 (1-93) x US-10-090-035-21 (1-436)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db      54  ATGGCGCACTTCAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGCGGTGGGCAACCCG 113
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPhe--- 38
Db      114  GCC-----CGCCGCGCGCGCGCGTGCAGGAGCAGCATCGTCAAGAGAGCGTCTGTG 164
QY      39  GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisHis 58
Db      165  CAGGAGTTCGACACCTCCGGCGCG-----CGCCAGCGTCACCGGTCCACCAC 212
QY      59  Gly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspPheAsn 75
Db      213  GGCGCGCGGTCTGTGTCTAC---TTCGAGGTGCGCGAGAGCAGGCTCGAGGAGACTTCAAC 269
QY      76  ThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      270  ACCCGCAGCGGGAGTTCACGAGCGCAAGGAGAACTTCGTGCTCAGGCGCGAT 323

RESULT 9
US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:
```

```
Db      279  GGCCACGGCTTCGTGTCGGCAGACAGGCTCGAGGAGGACATCAACACCTGCACCGGC 338
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      339  GAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGGCTAAC 380

RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.:      4e-30      Length:      591
Score:          295.00     Matches:      57
Percent Similarity: 75.53%  Conservative: 14
Best Local Similarity: 60.64% Mismatches:    19
Query Match:     57.84%   Indels:        4
DB:              9       Gaps:          2
US-10-090-035-2 (1-93) x US-10-090-035-15 (1-591)
QY      1  MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db      61  ATGGCTACTACCAAGAGGTGGACTACTGCTCGAGGAGGTGAGGTGGTGACCCACC 120
QY      20  AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db      121  GGCGGCTTCCTCGCGCGCGCGGCGTGCAGCAGCAGCAGCTCGTCAAGAGAGAGCTTCCAG 180
QY      40  GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db      181  GAGATCGAC-----AGTCCGGCTCCGGCGCCGCCACCAACCAACCAACCAACCAAC 231
QY      60  GlyHisGlyPheValArgGluThrArgValGluGluAspPheLeuAlaArgAlaAsn 79
Db      232  AACGACTACCTGATGGTGGCGGAGCAAGAGTGGAGGAGGAGTCAACACCTGCACCGGC 291
QY      80  GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db      292  GAGTTCGCGAGCGGAGCAGAGCTTCCTGCTCAAGTCCGAC 333

RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
```

Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-2 (1-93) x US-10-090-035-19 (1-584)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 46 ATGGCGCAGTTCAGGAGGTGGACTACTGCTCGGAGGAGTTCAGGCGGTGGCTACCG 105
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
Db 106 GCC-----CGCGCGGCTGCGCGCGGTGCGAGGACATCGTCAAGGAGAGCGTTC 156
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHis 57
Db 157 GTGAGGAGTTCAGACACCGCGCGCGC-----CGCCAYGGTCCACCGGTCCAC 204
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
Db 205 CACGCGCGGCGTCTGGTCAC---TTCGAGGTGCGGAGCAGCAAGCTSGARGAGCATC 261
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 262 AACACCCGACCGGSGAGTTCACGACGCAAGGGAAYTCTCTCCAGGCGGAT 318

RESULT 10
US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-23
Alignment Scores:
Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-2 (1-93) x US-10-090-035-23 (1-584)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 46 ATGGCGCAGTTCAGGAGGTGGACTACTGCTCGGAGGAGTTCAGGCGGTGGCTACCG 105
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
Db 106 GCC-----CGCGCGGCTGCGCGCGGTGCGAGGACATCGTCAAGGAGAGCGTTC 156

QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHis 57
Db 157 GTGAGGAGTTCAGACACCGCGCGCGC-----CGCCAYGGTCCACCGGTCCAC 204
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspIle 74
Db 205 CACGCGCGGCGTCTGGTCAC---TTCGAGGTGCGGAGCAGCAAGCTSGARGAGCATC 261
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 262 AACACCCGACCGGSGAGTTCACGACGCAAGGGAAYTCTCTCCAGGCGGAT 318

RESULT 11
US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc_feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-13
Alignment Scores:
Pred. No.: 1.39e-24 Length: 348
Score: 251.00 Matches: 51
Percent Similarity: 70.59% Conservative: 9
Best Local Similarity: 60.00% Mismatches: 21
Query Match: 49.22% Indels: 4
DB: 9 Gaps: 2

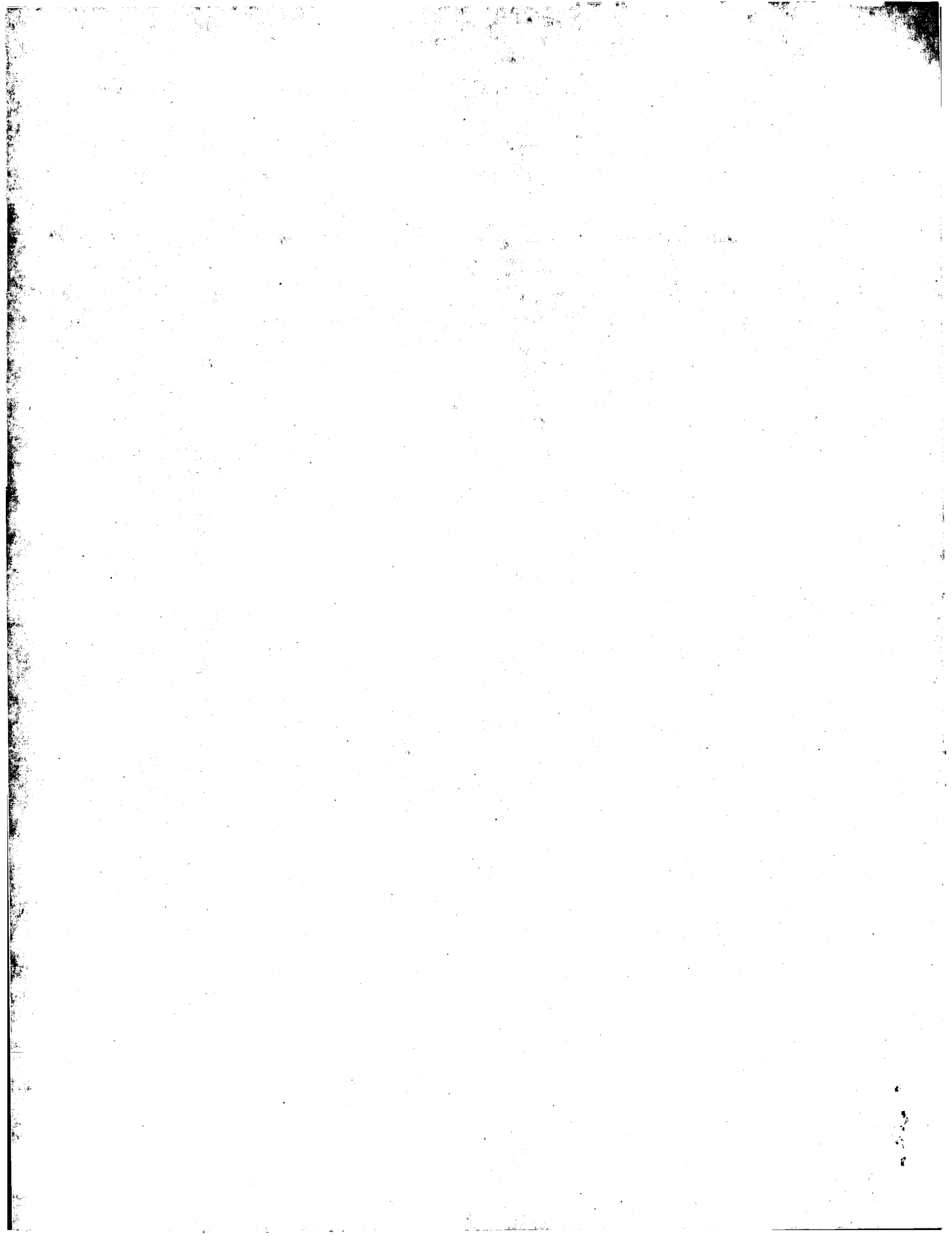
US-10-090-035-2 (1-93) x US-10-090-035-13 (1-348)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db 52 ATGGCTCAGTACAGGAGGTGGACTACTGCTCGGAGGAGTTCGCTGCGTACCCGCC 111
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db 112 GCGGCTTCTCGCGCGCGCGCGTTCGAGCAGCAGCAGCGTTCGTCAGGAGTTCAG 171
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHis 59
Db 172 GAGATCGACANG-----TCCGGCTCCGCGCGCGCAGCAGCAGCAGCAGCAGC 222
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 223 AACGACTACCTNATGTCGCGAGACCAAGGTGAGGAGGACTTTTAAACACCTGCACCG 282
QY 80 GluValHisGluArg 84
Db 283 GAGTTTCGCGAGCGC 297

RESULT 12
US-09-923-876-2788
; Sequence 2788, Application US/09923876

[illegible]

Search completed: June 16, 2003, 13:51:16
Job time : 132 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:48:33 ; Search time 25 Seconds
(without alignments)
357.620 Million cell updates/sec

Title: US-10-090-035-2
Perfect score: 510
Sequence: 1 MAYOEVDCSEEVRSVAPA.....INTCTGVEHRESFLARAN 93

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92.5	18.1	196	2 G85435	TINY-like protein
2	92.5	18.1	259	2 T52619	TINY-like protein
3	92	18.0	414	2 A48273	delta/YVI/NF-EI/UC
4	89	17.5	1245	2 T49815	related to multifu
5	86.5	17.0	471	2 T33997	hypothetical prote
6	84	16.5	473	2 D96591	hypothetical prote
7	82.5	16.2	735	2 T45059	hypothetical prote
8	82	16.1	173	2 T51469	glycine/proline-ri
9	81.5	16.0	495	1 S31223	transcription fact
10	81	15.9	86	2 T16437	hypothetical prote
11	81	15.9	191	2 A42844	absicisic acid-and
12	80	15.7	133	2 T09608	environmental stre
13	80	15.7	1273	2 T00338	hypothetical prote
14	79	15.5	204	2 T09592	protein corA, cold
15	79	15.5	414	2 A40350	transcription repr
16	77	15.1	754	2 JC4898	Down-syndrome-crit
17	76.5	15.0	77	2 T16436	hypothetical prote
18	76.5	15.0	332	2 C83682	hypothetical prote
19	76	14.9	102	2 T30119	hypothetical prote
20	76	14.9	1500	2 T00080	transcription fact
21	75.5	14.8	425	1 JH0710	transforming prote
22	75	14.7	369	1 TVEVAF	probable limonene
23	75	14.7	600	2 F71434	dve protein - frul
24	75	14.7	1019	2 T00117	CRAG protein - fru
25	75	14.7	1441	2 T13717	hypothetical prote
26	74.5	14.6	410	2 T26757	hypothetical prote
27	74	14.5	83	2 T16435	hypothetical prote
28	74	14.5	895	2 JC7089	zinc finger bindin
29	74	14.5	1891	2 T13594	hypothetical prote

30	73.5	14.4	147	2 T16440	hypothetical prote
31	73.5	14.4	285	2 E84766	probable Ar-hook D
32	73.5	14.4	499	2 S09880	hypothetical prote
33	73.5	14.4	776	2 T02702	hypothetical prote
34	73	14.3	606	2 A41145	gamma-aminobutyric
35	73	14.3	633	1 A26030	serine/threonine-s
36	72.5	14.2	1585	2 T31611	hypothetical prote
37	72	14.1	354	2 S39406	homeotic protein o
38	71.5	14.0	268	2 A56446	ig heavy chain v r
39	71.5	14.0	499	2 C75251	phosphoryruvate hy
40	71.5	14.0	699	2 A54660	histidine rich cal
41	71	13.9	338	2 A82890	hypothetical prote
42	71	13.9	361	2 F87286	cation efflux faml
43	71	13.9	2038	2 A43742	female sterile hom
44	71	13.9	2396	2 T13714	kakapo gene protei
45	70.5	13.8	59	2 T09610	environmental stre

ALIGNMENTS

RESULT 1

G85435

TINY-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: G85435

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: G85435

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <STO>

A:Cross-references: GB:NC_001268; NID:g7270639; PIDN:CAB80356.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g36900

A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.012;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY 21 GFGHGGG-VQOHVHKEFEE----VDTVSRAGA---NHHHHHGH-HGGHGFVVRER--R 69
Db 99 GGGVGGGDMSAAYTRKAAEVGAQVDALEAAGAGGNRRHHHHHHQHGRNHDYVDNHSYR 158

QY 70 VEEDINTCTGE 80
Db 159 INDDLMECSK 169

RESULT 2

T52619

TINY-like protein [imported] - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000

C:Accession: T52619

R:terryn, N.; Heijnen, L.; De Keyser, A.; Van Asseldonck, M.; De Clercq, R.; Verbakel

ueller, C.; Mayer, K.; Dehais, P.; Rombauts, S.; Van Montagu, M.; Rouze, P.; Vos, P.

FEBS Lett. 445, 237-245, 1999

A:Title: Evidence for an ancient chromosomal duplication in Arabidopsis thaliana by

A:Reference number: Z26022; MUID:99192287; PMID:10094464

A:Accession: T52619

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-259 <TER>

A:Cross-references: EMBL:AJ002598; PIDN:CAA05630.1

C:Genetics:

A:Map position: 4

Query Match 18.1%; Score 92.5; DB 2; Length 259;

Mon Jun 16 14:55:55 2003

```
Best Local Similarity 33.8%; Pred. No. 0.016;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY 21 GFGHGGG-VQOHVVKFEE----VDTVSRAGA---NHHHHHGH-HGCHGVVRET--R 69
Db 162 GGVNGGDMGSAYTRRAAEVGAQVDALAAAGAGNRHHHHQHQRGHHHDYVDNHSYR 221
QY 70 VEEDINTCTGE 80
Db 222 INDDLMECSK 232

RESULT 3
A48273
delta/VYI/NF-El/UCRBP transcription factor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C:Accession: A48273; A42055; A56418
R:Safrany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A:Title: Characterization of the mouse gene that encodes the delta/VYI/NF-El/UCRBP trans
A:Reference number: A48273; MUID:93296177; PMID:8516301
A:Accession: A48273
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-414 <RES>
A:Cross-references: GB:I13968; NID:9293847; PIDN:AAA40477.1; PID:9293849
R:Flanagan, J.R.; Becker, K.G.; Ennlist, D.L.; Gleason, S.L.; Driggers, P.H.; Levi, B.Z.;
Mol. Cell. Biol. 12, 38-44, 1992
A:Title: Cloning of a negative transcription factor that binds to the upstream conserved
A:Reference number: A42055; MUID:92107191; PMID:1309593
A:Accession: A42055
A:Molecule type: mRNA
A:Residues: 1-414 <FLA>
A:Cross-references: GB:M73963; NID:g202270; PIDN:AAA40522.1; PID:g202271
A:Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:P:74642)
R:Harisharan, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A:Title: Delta, a transcription factor that binds to downstream elements in several poly
A:Reference number: A56418; MUID:92052178; PMID:1946404
A:Accession: A56418
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218, 'S', 220-374, 'G', 376-414 <HAR>
A:Cross-references: GB:M74590; NID:g192940; PIDN:AAA37521.1; PID:g192941
C:Genetics:
A:Introns: 227/1; 281/2; 301/3; 354/3
C:Keywords: transcription factor; zinc finger

Query Match 18.0%; Score 92; DB 2; Length 414;
Best Local Similarity 31.0%; Pred. No. 0.031;
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps 2;

QY 21 GFGHGGGVQOHVVKFEEVDIVSRAGANHHHHHGHGGHGVFWRETRVEEDINTCTGE 80
Db 55 GGGDHGGGGGGH-----GHAGHHHHHHHHHHPPMIALQPLVTDD----PTQ 98
QY 81 VHERRESFLAR 91
Db 99 VHHQEVILVQ 109

RESULT 4
T49815
related to multifunctional cyclin-dependent kinase PH085 [imported] - Neurospora crassa
N:Alternate names: protein B24H17.10
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
A:Accession: T49815
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49815
```

```
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1245 <SCH>
A:Cross-references: EMBL:AJ356815; GSPDB:GN00116; NCSP:B24H17.10
A:Experimental source: BAC clone B24H17; strain OR74A
C:Genetics:
A:Gene: NCSP:B24H17.10
A:Map position: 6
A:Introns: 6/2; 1141/3

Query Match 17.5%; Score 89; DB 2; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.21;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 45 SRAGANHHHHHGHGGHG 62
Db 901 SRSGTHHHHHHHHHGHG 918

RESULT 5
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33997
R:Pauley, A.; Scheet, P.; Harper, M.
Submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: 221454
A:Accession: T33997
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: EMBL:AF125964; PIDN:AAD14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 17.0%; Score 86.5; DB 2; Length 471;
Best Local Similarity 34.8%; Pred. No. 0.14;
Matches 24; Conservative 4; Mismatches 16; Indels 25; Gaps 4;

QY 15 RSVAPAGFRHGG-----GVQOHVVKFEEVDIVSRAGANHHHHHGHG---- 56
Db 372 RSRSPRGH-HGRRHGPCHPCORHGHGPPH-----HHHDGRSPSRHHHHHHHGGCRPF 426
QY 57 --HHGGHGF 63
Db 427 PPHGHGHHF 435

RESULT 6
D96591
hypothetical protein T24C10.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96591
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maity, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <STO>
```

A:Cross-references: GB:AE005173; NID:99857523; PIDN:AAG00878.1; GSPDB:GN00141
C:Genetics:
A:Gene: T24C10.10
A:Map position: 1

Query Match 16.5%; Score 84; DB 2; Length 473;
Best Local Similarity 36.8%; Pred. No. 0.26; Mismatches 18; Indels 8; Gaps 3;
Matches 21; Conservative 10;

QY 8 DYCEEVRSVAPAGFGHGGVQOHHVVKKEFEVDVTSRAGANHHHHGHG---HGCG 61
DB 409 DYISEFV-SLLPKSIRR-----VAEEPIPEVQVLEAKAGDDHDDHHGHGHAHAGY 460

RESULT 7

T45059
hypothetical protein Y39B68.gg [Imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
R:Wilson, R.; Alnscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.
Nature 368, 32-38, 1994
A:Authors: Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stock, L.; Wilkinson-Sproat, J.; Wooldman, P.
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID:94150718; PMID:7906398
A:Accession: T45059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-735 <WIL>
A:Cross-references: EMBL:AL128896; NID:96434440; PIDN:CAB60938.1; PID:96434473
A:Experimental source: clone Y39B6B
C:Genetics:
A:Map position: 3
A:Introns: 18/1: 69/1
A:Note: Y39B6B.gg

Query Match 16.2%; Score 82.5; DB 2; Length 735;
Best Local Similarity 26.4%; Pred. No. 0.6; Mismatches 19; Conservative 1; Indels 43; Gaps 2;
Matches 19;

QY 23 GRHGGVQOHHVVKKEFEVDVTSRAGANHHH-----HGHHGAHHHAPHHEHHRHGHGSHGVHHGHGTHH 662
DB 614 GHGAGYGAN-----HGHGAGYGAN-----53
QY 54 ---HHHGHGGH 62
DB 663 SLAHHGHGGH 674

RESULT 8

T51469
glycine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein K10A8_130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
R:Sato, S.; Nakamura, Y.; Kaneo, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <SAT>
A:Cross-references: EMBL:AL391151
A:Experimental source: cultivar Columbia; BAC clone K10A8
C:Genetics:
A:Map position: 5
A:Introns: 97/1
A:Note: K10A8_130

Query Match 16.1%; Score 82; DB 2; Length 173;
Best Local Similarity 34.7%; Pred. No. 0.15; Mismatches 17; Conservative 7; Indels 10; Gaps 2;
Matches 17;

QY 19 PAGFGRHGGVQOHHVVKKEFEVDVTSRAGA-----NHHHHHHHHGGHGF 63
DB 93 PSHSHHHGGIGAILAG-----GVAAAGAAHHSHHHHHGHGHHGHGY 135

RESULT 9

S31223
transcription factor Brn-1 - mouse
N:Alternate names: class III POU domain protein brain-1
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
C:Accession: S31223
R:Hara, Y.; Rovescalli, A.C.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A:Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A:Reference number: S31223; MUID:92228768; PMID:1565620
A:Accession: S31223
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAR>
A:Cross-references: EMBL:M89299; NID:9200444; PIDN:AAA39960.1; PID:9200445
C:Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:27-49/Region: glycine-rich
F:101-112/Region: alanine-rich
F:162-180/Region: histidine/proline-rich
F:186-201/Region: alanine-rich
F:236-247/Region: glycine-rich
F:267-291/Region: histidine/proline-rich
F:316-383/Domain: POU domain homology <POU>
F:402-458/Domain: homeobox homology <HOX>

Query Match 16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.5; Mismatches 24; Conservative 7; Indels 25; Gaps 2;
Matches 24;

QY 19 PAGFGRHGGVQOHHVVKKEFEVDVTSRAGANHHHH-----GHHGGHG- 62
DB 237 PGCGGGAGGAGSLVHPLVGRDTPELAEHHHHHHHHHAPHPPHHAQGPPIHGGGA 296
QY 63 -----FVRETRVEEDINTCTGEVHERR 85
DB 297 GPGLNSHDPHSDDEDTPTSDLEQFAKQFKORR 328

RESULT 10

T16437
hypothetical protein F53A9.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16437
R:Miller, N.
submitted to the EMBL Data Library, March 1995
A:Description: The sequence of C. elegans cosmid F53A9.
A:Reference number: Z19513
A:Accession: T16437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-86 <MIL>
A:Cross-references: EMBL:U23523; NID:9746551; PID:9746557; PIDN:AAC46561.1; CBSP:F53A
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CBSP:F53A9.6

Query Match 15.9%; Score 81; DB 2; Length 86;
Best Local Similarity 29.5%; Pred. No. 0.089; Mismatches 23; Conservative 5; Indels 24; Gaps 3;
Matches 23;

QY 1 MAYQEVDCSEVRSVAPAGFGRHGG-----GVQOHVVKKEFEVDVTS 45

C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00338
C:R:R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545; PMID:9628581
A:Accession: T00338
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <NAG>
A:Cross-references: EMBL:AB011142; NID:g3043663; PIDN:BAA25496.1; PID:g3043664
A:Experimental source: brain; clone HH2365
C:Genetics:
A:Note: KIAA0570 human cytothetical protein KIAA0570

	Query Match	15.7%	Score 80;	DB 2;	Length 1273;	
	Best Local Similarity	33.3%	Pred. No. 2;			
	Matches 15;	Conservative	9;	Mismatches	19;	Indels
						Gaps
QY	48	GANHHHGHGHH-	-GFVVRETRVEEDINTCTGEVHERRESFLA	90		
			: : : : : : : :			
Dd	595	FQGHHHHHHHHHHGHHGVDDMLSADDVSCSSQVSAKSEKNMA	639			

RESULT 14

A:Residues: 1-204 <LAB>
A:Cross-references: EMBL:L03708; NID:g289122; PIDN:AAA99833.1; PID:g289123
C:Genetics:
C:Gene: corA
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match	15.5%	Score 79;	DB 2;	Length 204;
Best Local Similarity	30.4%	Pred. No. 0.37;		
Matches	24;	Conservative	24;	Indels 28; Gaps 3;

RESULT 15

A;Molecule type: mRNA
A;Residues: 1-414 <SH1>
A;Cross-references: NC_077698: NID: g186767: PIDN: AAS59467.1: PID: g186768

A:Experimental source: HeLa cells
A:Note: the authors translated the codon CGC for residue 371 as Lys, CGA for residue 375
R:Whitson, R.H.; Huang, T.; Pang, J.; Itakura, K.
submitted to the EMBL Data Library, July 1992
A:Description: Observed and predicted DNA binding of a zinc finger protein which recogni
A:Reference number: S78494
A:Accession: S78494
A:Molecule type: mRNA
A:Residues: 1-195,'G',197-414 <WHI>
A:Cross-references: EMBL:J14077; NID:g38010; PIDN:CAA78455.1; PID:g38011
R:Park, K.; Atchison, M.L.
Proc. Natl. Acad. Sci. U.S.A. 88, 9804-9808, 1991
A:Title: Isolation of a candidate repressor/activator, NF-E1 (YY-1), delta, that binds t
A:Reference number: A56419; MUID:92052179; PMID:1946405
A:Accession: S33712
A:Molecule type: mRNA
A:Residues: 1-64,'R',66-195,'G',197-414 <PAR>
A:Cross-references: GB:M76541; NID:gl89173; PIDN:AAA59926.1; PID:gl89174
C:Keywords: DNA binding; transcription regulation; zinc finger
F:298-320/Region: zinc finger CCHH motif
F:327-347/Region: zinc finger CCHH motif
F:355-377/Region: zinc finger CCHH motif
F:385-407/Region: zinc finger CCHH motif

Query Match	15.5%	Score 79;	DB 2;	Length 414;
Best Local Similarity	29.2%;	Pred. No. 0.78;		
Matches	21;	Conservative	6;	Mismatches 25; Indels 20; Gaps 3;
QY	21	GGRIGGGVQOHVWKEFEEVDTVSRAGANHHHHHHGGH	-GFVVRTVEEDINTCTG	79
Db	55	GGGDHGGG-----	GGHGAGHHHHHHHHHPPMIALQLVTTDD	-----PT 95
QY	80	EVHERRESFLAR	91	
		: : :	: :	
Db	96	QVHHHQEVILQ	107	

Search completed: June 6, 2003, 12:53:24
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:42:32 ; Search time 13 Seconds
(without alignments)
296.715 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYQEVDCSEVRVAP.....INTCTGEVHERSEFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	18.0	414	1	TYYL_MOUSE
2	81.5	16.0	495	1	BRNL_MOUSE
3	81.5	16.0	497	1	BRNL_RAT
4	80.5	15.8	500	1	BRNL_HUMAN
5	79	15.5	204	1	CORA_MEDSA
6	79	15.5	414	1	TYYL_HUMAN
7	77	15.1	763	1	DYRA_HUMAN
8	77	15.1	763	1	DYRA_MOUSE
9	77	15.1	763	1	DYRA_RAT
10	76	14.9	403	1	MAF_HUMAN
11	76	14.9	551	1	CBX4_MOUSE
12	75.5	14.8	425	1	POU1_BRARE
13	75.5	14.8	558	1	CBX4_HUMAN
14	75	14.7	369	1	TMAF_AVIS4
15	74	14.5	449	1	CSUP_DROME
16	74	14.5	895	1	Z281_HUMAN
17	74	14.5	2190	1	CCAD_CHICK
18	73.5	14.4	147	1	YV59_CAEEL
19	73.5	14.4	684	1	EP84_HCMVA
20	73	14.3	606	1	GAB_DROME
21	73	14.3	633	1	SNF1_YEAST
22	72	14.1	354	1	OTX1_HUMAN
23	71.5	14.0	459	1	A2AC_DIDMA
24	71.5	14.0	532	1	ZIC2_HUMAN
25	71.5	14.0	699	1	SRCH_HUMAN
26	71	13.9	2038	1	FSH_DROME
27	70.5	13.8	139	1	SALA_DROSI
28	70.5	13.8	355	1	OTX1_MOUSE
29	70.5	13.8	355	1	OTX1_RAT
30	70.5	13.8	727	1	CCRL_HORSE
31	70.5	13.8	2175	1	HMCU_DROME
32	70	13.7	190	1	HUNB_DROME
33	70	13.7	370	1	MAF_MOUSE

34	69.5	13.6	515	1	KE4L_CAEEL
35	69.5	13.6	623	1	PNT1_DROME
36	69.5	13.6	726	1	CCT1_HUMAN
37	69	13.5	85	1	ANTF_SARPE
38	69	13.5	332	1	DLX2_MOUSE
39	69	13.5	351	1	CAV2_CAEEL
40	69	13.5	359	1	HNK1_MAIZE
41	69	13.5	410	1	BR3B_HUMAN
42	69	13.5	411	1	BR3B_MOUSE
43	68.5	13.4	323	1	MAFB_HUMAN
44	68.5	13.4	873	1	RX_DROME
45	68	13.3	351	1	HRPX_PLALO

ALIGNMENTS

RESULT 1
TYYL_MOUSE
ID TYYL_MOUSE STANDARD; PRT; 414 AA.
AC Q00899;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE transcriptional repressor protein YY1 (Yin and yang 1) (YY-1) (Delta
transcription factor) (NF-E1) (UCR-motif DNA-binding protein).
GN YY1 OR UCRBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92107191; PubMed=1309593;
RA Flanagan J.R., Becker K.G., Ennist D.L., Gleason S.L., Driggers P.H.,
RA Levi B.-Z., Appella E., Ozato K.;
RT "Cloning of a negative transcription factor that binds to the
upstream conserved region of Moloney murine leukemia virus";
RL Mol. Cell. Biol. 12:38-44(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93296177; PubMed=8516301;
RA Safrany G., Perry R.P.;
RT "Characterization of the mouse gene that encodes the delta/YY1/NF-
El/UCRBP transcription factor";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5559-5563(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92052178; PubMed=1946404;
RA Hariharan N., Kelley D.E., Perry R.P.;
RT "Delta, a transcription factor that binds to downstream elements in
several polymerase II promoters, is a functionally versatile zinc
finger protein";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9799-9803(1991).
CC -!- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT AND
DIFFERENTIATION. THE FUNCTION OF YY1 AS AN ACTIVATOR OR A
REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. BINDS TO
THE UPSTREAM CONSERVED REGION (UCR) (5'- GCCTATT-3') OF MOLONEY
MURINE LEUKEMIA VIRUS (MULV).
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@sib-sib.ch).

EMBL: M73963; AAA40522.1; -
EMBL: L13968; AAA40477.1; -
EMBL: L13969; AAA40477.1; JOINED.
EMBL: L13965; AAA40477.1; JOINED.
EMBL: L13966; AAA40477.1; JOINED.
EMBL: L13967; AAA40477.1; JOINED.
EMBL: M74590; AAA37521.1; -
HSP: P25490; IUBD.
TRANSFAC: T00278; -
TRANSFAC: T00865; -
MGD; MGI:99150; Y1.
InterPro: IPR000822; Znf_C2H2.
PRINTS; PR00096; zf-C2H2; 4.
PFAM; PF00048; ZINC_FINGER.
PRODOM; PD000003; Znf_C2H2; 1.
SMART; SM00355; Znf_C2H2; 4.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
Transcription regulation; Repressor; Activator; Nuclear protein;
Zinc-finger; Metal-binding; DNA-binding; Repeat.
DOMAIN 43 53 ASP/GLU-RICH (ACIDIC).
DOMAIN 54 70 GLY-RICH.
DOMAIN 71 82 POLY-HIS.
DOMAIN 161 170 GLY/SER-RICH.
DOMAIN 296 407 ZINC_FINGERS.
ZNF_FING 296 407 ZINC_FINGERS.
ZNF_FING 325 347 C2H2-TYPE.
ZNF_FING 353 377 C2H2-TYPE.
ZNF_FING 383 407 C2H2-TYPE.
DOMAIN 257 341 INVOLVED IN NUCLEAR MATRIX ASSOCIATION
(BY SIMILARITY).
DOMAIN 333 371 INVOLVED IN REPRESSION OF ACTIVATED
TRANSCRIPTION (BY SIMILARITY).
DOMAIN 371 397 INVOLVED IN MASKING TRANSCRIPTION
DOMAIN (BY SIMILARITY).
CONFLICT 219 219 F -> S (IN REF. 3).
CONFLICT 375 375 R -> G (IN REF. 3).
SEQUENCE 414 AA; 44717 MW; C01237828E984F9 CRC64;

Query Match 18.0%; Score 92; DB 1; Length 414;
Best Local Similarity 31.0%; Pred. No. 0.017;
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps 2;

QY 21 GFGHGGGVQVHVVKKEFEVDVTSRAGANHHHHHGGHGGFVVRETRVEEDINTCTGE 80
DB 55 GGDHGGGGGGH-----GHAGHHHHHHHHHHHHHPMTALQPLVTD-----PTQ 98

QY 81 VHERRESFLAR 91
DB 99 VHHQEVILVQ 109

RESULT 2
BRNL_MOUSE
ID BRNL_MOUSE STANDARD; PRT; 495 AA.
AC P31361;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).
GN POU3F3 OR OTF8 OR BRN1 OR BRN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92228768; PubMed=1565620;
RA Hara Y., Rovescalli C., Kim Y., Nirenberg M.;
RT "Structure and evolution of four POU domain genes expressed in mouse
brain.";

Proc. Natl. Acad. Sci. U.S.A. 89:3280-3284(1992).

-1- SUBCELLULAR LOCATION: Nuclear.
-1- TISSUE SPECIFICITY: BRAIN.
-1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CLASS-3 SUBFAMILY.
-1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: M88299; AAA39960.1; -
PIR: S31223; S31223.
HSP; P14859; 10CT.
MGD; MGI:102564; Pou3f3.
InterPro: IPR001356; Homeobox.
InterPro: IPR000327; POU_domain.
Pfam: PF00046; homeobox; 1.
Pfam: PF00157; pou; 1.
PRINTS; PR00028; POU_DOMAIN.
PRODOM; PD000010; Homeobox; 1.
PRODOM; PD000583; POU_domain; 1.
SMART; SM00389; HOX; 1.
SMART; SM00352; POU; 1.
PROSITE; PS00027; HOMEBOX_1; 1.
PROSITE; PS00071; HOMEBOX_2; 1.
PROSITE; PS00035; POU_1; 1.
PROSITE; PS00465; POU_2; 1.
Nuclear protein; DNA-binding; Homeobox.
KW DOMAIN 28 49 POLY-GLY.
FT DOMAIN 101 112 POLY-ALA.
FT DOMAIN 186 201 POLY-ALA.
FT DOMAIN 267 291 HIS-RICH.
FT DOMAIN 313 383 POU.
FT DNA_BIND 401 460 HOMEBOX.
FT SEQUENCE 495 AA; 50012 MW; 778802E890C9A014 CRC64;

Query Match 16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.26;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;

QY 19 PAGFGHGGGVQVHVVKKEFEVDVTSRAGANHHHH-----GHHGGH- 62
DB 237 PGGGGGGAGGAGQSLVHPGLVGRDTPPELAHHHHHHHHHHHHPPHPPHHAQGP HHGGGA 296

QY 63 -----FVYRETRVEEDINTCTGEVHERR 85
DB 297 GPCGLNSHDPHSDPTSDLEQFAKQKRR 328

RESULT 3
BRNL_RAT
ID BRNL_RAT STANDARD; PRT; 497 AA.
AC Q63262;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).
GN POU3F3 OR BRN1 OR BRN-1 OR RNS2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98070400; PubMed=9405434;
RA Schreiber J., Enderich J., Sock E., Schmidt C., Richter-Landsberg C.,
Wegner M.;
RT "Redundancy of class III POU proteins in the oligodendrocyte

```
RT lineage.";
RN J. Biol. Chem. 272:32286-32293(1997).
RP [2]
RC SEQUENCE OF 325-449 FROM N.A.
RX TISSUE-Hypothalamus;
MEDLINE=92228769; PubMed=1348858;
le Moine C., Young W.S.;
"RHS2, a POU domain-containing gene, and its expression in developing
and adult rat.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:3285-3289(1992).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM EMBRYONIC DAY 11.5 INTO
CC ADULTHOOD.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-3 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001641; CA04893.1; -
DR EMBL; M84644; AAA42041.1; -
DR HSSP; P14859; LOCT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00157; pou; 1.
DR PRODOM; PD000028; POU DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00583; POU_domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 28 49 POLY-GLY.
FT DOMAIN 103 114 POLY-ALA.
FT DOMAIN 135 143 POLY-PRO.
FT DOMAIN 173 178 POLY-PRO.
FT DOMAIN 188 203 POLY-ALA.
FT DOMAIN 238 249 POLY-GLY.
FT DOMAIN 269 280 POLY-HIS.
FT DOMAIN 294 301 POLY-GLY.
FT DOMAIN 315 385 POU.
FT DNA_BIND 403 462 HOMEBOX.
SQ SEQUENCE 497 AA; 50226 MW; 00640505E343ABC2 CRC64;

Query Match
Best Local Similarity 16.0%; Score 81.5; DB 1; Length 497;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;

QY 19 PAGFGRRGGGQQHVYKEFEVDVTSRAGANHHHH-----GHHGGHG- 62
Db 239 PGGGGGGAGGAGSLVHPGLVGRDTPELAEHHHHHHHHHPHPHPHQAQPPHGGGA 298
QY 63 -----FVRETRVEDINTCTGEVHER 85
Db 299 GPGLSHDPSHDEDTPTSDLEQFAQFKQRR 330

RESULT 4
BRN1_HUMAN
ID P20264; P78379; STANDARD; PRT; 500 AA.
DT 01-FEB-1991 (Rel. 17, Created)
```

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brain-specific homeobox/POU domain protein 1 (BRN-1 protein).
GN POU3F3 OR BRN1 OR OTF8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96359175; PubMed=8703082;
RA Suniyama K., Washio-Watanabe K., Saitou N., Hayakawa T., Ueda S.;
RT "Class III POU genes: generation of homopolymeric amino acid repeats
RT under GC pressure in mammals.";
RL J. Mol. Evol. 43:170-178(1996).
RN [2]
RP SEQUENCE OF 332-456 FROM N.A.
RC TISSUE-Brain.
RX MEDLINE=89295573; PubMed=2739723;
RA He X., Treacy M.N., Simmons D.M., Ingraham H.A., Swanson L.W.,
RA Rosenfeld M.G.;
RT "Expression of a large family of POU-domain regulatory genes in
RT mammalian brain development.";
RL Nature 340:35-42(1989).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-3 SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB001835; BAA19459.1; -
DR FIR; S05042; S05042.
DR HSSP; P14859; LOCT.
DR TRANSFAC; T04469; -
DR Genew; HGNC:9216; POU3F3.
DR MIM; 602480; -
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000327; POU_domain.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00157; pou; 1.
DR PRINTS; PR00028; POU DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU_domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 28 50 POLY-GLY.
FT DOMAIN 104 115 POLY-ALA.
FT DOMAIN 136 144 POLY-PRO.
FT DOMAIN 174 179 POLY-PRO.
FT DOMAIN 189 204 POLY-ALA.
FT DOMAIN 239 250 POLY-GLY.
FT DOMAIN 270 294 POLY-HIS.
FT DOMAIN 295 304 POLY-GLY.
FT DOMAIN 318 388 POU.
FT DNA_BIND 406 465 HOMEBOX.
FT CONFLICT 433 433 A -> S (IN REF. 2).
SQ SEQUENCE 500 AA; 50327 MW; E536EFFFA5212319 CRC64;

Query Match
15.8%; Score 80.5; DB 1; Length 500;
```

Best Local Similarity 33.9%; Pred. No. 0.34;
Matches 20; Conservative 2; Mismatches 22; Indels 15; Gaps 1;
QY 19 PAFGRHGGVQOHHVKEFEVDVSVRAGANHHHH-----GHHGGH 62
DB 240 PGGGGGAGGAQSLVHFLVGRDTPELAHHHHHHHHHAHPHPHPHHAQPPHHGGG 298

RESULT 5
CORAMEDSA STANDARD; PRT; 204 AA.
ID CORA_MEDSA
AC Q07202; 1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Cold and drought-regulated protein CORA.
GN CORA.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbiales I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Adica;
RX MEDLINE=94143496; PubMed=8310076;
RA Laberge S., Castonguay Y., Vezina L.-P.;
RT "New cold- and drought-regulated gene from Medicago sativa.";
RL Plant Physiol. 101:1411-1412(1993).
CC -1- FUNCTION: MAY BE INVOLVED IN RESISTANCE OF THE PLANT TO
ENVIRONMENTAL STRESS.
CC -1- INDUCTION: BY COLD, ABSCISIC ACID (ABA) AND DROUGHT STRESS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L03708; AAA9833.1;
KW Multigene family; Repeat.
FT DOMAIN 54 176 7 X 6 AA REPEATS OF Y-N-H-G-G-G.
FT REPEAT 54 59 1-1.
FT REPEAT 65 70 1-2.
FT REPEAT 71 76 1-3.
FT REPEAT 78 83 1-4.
FT REPEAT 85 90 1-5.
FT REPEAT 164 169 1-6.
FT REPEAT 171 176 1-7.
FT DOMAIN 98 192 11 X 3 AA REPEATS OF H-G-G.
FT REPEAT 98 100 2-1.
FT REPEAT 101 103 2-2.
FT REPEAT 112 114 2-3.
FT REPEAT 115 117 2-4.
FT REPEAT 126 128 2-5.
FT REPEAT 129 131 2-6.
FT REPEAT 178 180 2-7.
FT REPEAT 181 183 2-8.
FT REPEAT 184 186 2-9.
FT REPEAT 187 189 2-10.
FT REPEAT 190 192 2-11.
SQ SEQUENCE 204 AA; 19599 MW; 8A2C082359FCC17F CRC64;

Query Match 15.5%; Score 79; DB 1; Length 204;
Best Local Similarity 30.4%; Pred. No. 0.19;
Matches 24; Conservative 3; Mismatches 24; Indels 28; Gaps 3;
QY 21 GFRHGGVQOHHV-KVEFEVDV-----SRAGAH-----H 52
DB 125 GHGGGGAESVAVQTEKTNEVDKAGGGSYNDGRGYNHGGGYNHGGGGHGGH 184

QY 53 HHHGGHGGHGFVVRETRVE 71
DB 185 GGHHGGHGGHGAQVQEDNTQ 203

RESULT 6
TYV1_HUMAN STANDARD; PRT; 414 AA.
ID TYV1_HUMAN
AC P25490; O14935;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcriptional repressor protein YV1 (yin and yang 1) (YY-1) (Delta
transcription factor) (NF-E1).
GN YV1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92005716; PubMed=1655281;
RA Shi Y., Seto E., Chang L.-S., Shenk T.;
RT "Transcriptional repression by YV1, a human GLI-Kruppel-related
protein, and relief of repression by adenovirus E1A protein.";
RL Cell 67:377-388(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Foreskin;
RX MEDLINE=92052179; PubMed=1946405;
RA Park K., Atchison M.;
RT "Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta),
that binds to the immunoglobulin kappa 3' enhancer and the
immunoglobulin heavy-chain mu E1 site.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9804-9808(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Whiskon R.H., Huang T., Dang J., Itakura K.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SUBCELLULAR LOCATION.
RX MEDLINE=98152931; PubMed=9493912;
RA McNeil S., Guo B., Stein J.L., Lian J.B., Bushmeyer S., Seto E.,
RA Atchison M.L., Penman S., van Wijnen A.J., Stein G.S.;
RT "Targeting of the YV1 transcription factor to the nucleolus and the
nuclear matrix in situ: the C-terminus is a principal determinant for
nuclear trafficking.";
RL J. Cell. Biochem. 68:500-510(1998).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 304-414.
RX MEDLINE=97098436; PubMed=8942976;
RA Houbaviy H.B., Usheva A., Shenk T., Burley S.K.;
RT "Cocystal structure of YV1 bound to the adeno-associated virus P5
initiator.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13577-13582(1996).
RN [6]
RP STRUCTURE BY NMR OF 353-379.
RX MEDLINE=98308000; PubMed=9642075;
RA Viles J.H., Patel S.U., Mitchell J.B.O., Moody C.M., Justice D.E.,
RA Uppenbrink J., Doyle P.M., Harris C.J., Sadler P.J., Thornton J.M.;
RT "Design, synthesis and structure of a zinc finger with an artificial
beta-turn.";
RL J. Mol. Biol. 279:973-986(1998).
CC -1- FUNCTION: MULTIFUNCTIONAL TRANSCRIPTION FACTOR THAT EXHIBITS
POSITIVE AND NEGATIVE CONTROL ON A LARGE NUMBER OF CELLULAR AND
VIRAL GENES BY BINDING TO SITES OVERLAPPING THE TRANSCRIPTION
START SITE. MAY PLAY AN IMPORTANT ROLE IN DEVELOPMENT OR A
DIFFERENTIATION. THE FUNCTION OF YV1 AS AN ACTIVATOR OR A
REPRESSOR IS SPECIFIED BY THE PRESENCE OF OTHER PROTEINS. FOR
EXAMPLE IT ACTS AS A REPRESSOR IN ABSENCE OF ADENOVIRUS E1A
PROTEIN BUT AS AN ACTIVATOR IN ITS PRESENCE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-


```
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
QY 11 SEEVSVAPAGFGHGG-----GVQHVVKKEFEVDTVSRA-----GANHHHHHG 56
DB 555 SPQVROQFPAPLGSFGTAPTVTVTHPQVETTFHVAPOQNALHHHGNSSHHHHHHH 614
QY 57 HHGHHG 62
DB 615 HHHHHG 620

RESULT 9
DYRA_RAT
ID DYRA_RAT STANDARD; PRT; 763 AA.
AC Q63470;
DT 15-JUL-1999 (Rel. 38, Last Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Dual-specificity tyrosine-phosphorylation regulated kinase 1A
DE (EC 2.7.1.1.) (Protein kinase minibrain homolog) (MNBH) (RP86) (Dual
DE specificity YAK1-related kinase).
GN DYRK1A OR DYRK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF TYR-219.
RC STRAIN-Sprague-Dawley; TISSUE-Brain;
RX MEDLINE=96216443; PubMed=8631952;
RA Kentrup H., Becker W., Heukelbach J., Wilmes A., Schuermann A.,
RA Huppertz C., Kainulainen H., Joost H.-G.;
RT "Dyrk, a dual specificity protein kinase with unique structural
RT features whose activity is dependent on tyrosine residues between
RT subdomains VII and VIII.";
RL J. Biol. Chem. 271:3488-3495(1996).
RN [2]
RP REVISIONS.
RA Kentrup H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98421512; PubMed=9748265;
RA Becker W., Weber Y., Wetzel K., Eilmbter K., Tejedor F.J.,
RA Joost H.-G.;
RT "Sequence characteristics, subcellular localization, and substrate
RT specificity of DYRK-related kinases, a novel family of dual
RT specificity protein kinases.";
RL J. Biol. Chem. 273:25893-25902(1998).
CC -!- FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
CC NUCLEAR FUNCTIONS OF CELL PROLIFERATION. PHOSPHORYLATES SERINES,
CC THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
CC SUBSTRATES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: UBQUITOUS.
CC -!- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MNB/DYRK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79769; CAA56164.1; -
CC HSP; Q00534; IBI8.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
```

```
Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing.
FT DOMAIN 117 134
FT BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT (POTENTIAL).
FT DOMAIN 159 479
FT NP_BIND 165 173
FT BINDING 188 188
FT ACT_SITE 287 287
FT BY_SIMILARITY.
FT DOMAIN 509 515
FT POLY-SER.
FT DOMAIN 599 602
FT POLY-HIS.
FT DOMAIN 607 619
FT SER/THR-RICH.
FT DOMAIN 656 672
FT MOD_RES 219 219
FT MOD_RES 319 319
FT MOD_RES 321 321
FT VARSPIC 70 78
FT MUTAGEN 219 219
FT Y->F: REDUCED AUTOPHOSPHORYLATION ON
FT TYROSINE, BUT NO LOSS OF HISTONE
FT PHOSPHORYLATION.
SQ SEQUENCE 763 AA; 85541 MW; CB5EC7EC4C1F9A47 CRC64;
Query Match 15.18; Score 77; DB 1; Length 763;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;
QY 11 SEEVSVAPAGFGHGG-----GVQHVVKKEFEVDTVSRA-----GANHHHHHG 56
DB 555 SPQVROQFPAPLGSFGTAPTVTVTHPQVETTFHVAPOQNALHHHGNSSHHHHHHH 614
QY 57 HHGHHG 62
DB 615 HHHHHG 620

RESULT 10
MAF_HUMAN
ID MAF_HUMAN STANDARD; PRT; 403 AA.
AC Q75444; Q9UP93;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transcription factor Maf (Proto-oncogene c-maf).
GN MAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98282186; PubMed=9616139;
RA Chesi M., Bersagel P.L., Shonukan O.O., Martelli M.L., Brents L.A.,
RA Chen T., Schrock E., Ried T., Kuehl W.M.;
RT "Frequent dysregulation of the c-maf proto-oncogene at 16q23 by
RT translocation to an Ig locus in multiple myeloma.";
RL Blood 91:4457-4463(1998).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -!- DISEASE: Some forms of multiple myeloma (MM) tumors are
CC characterized by a chromosomal translocation t(14;16)(q32.3;q23)
CC that involves MAF and an Igh locus.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. MAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; U63387; AAB96874.1; --
DR HSP; P23197; LAP0
DR MGD; MGI:1195985; Cbx4.
DR InterPro; IPR000953; Chromo.
DR Pfam; PF00385; chromo; 1.
DR PRINTS; PR00504; CHROMODOMAIN.
DR SMART; SM00298; CHROMO.1.
DR PROSITE; PS00598; CHROMO.1; 1.
DR PROSITE; PS00113; CHROMO.2; 1.
DR Chromatin regulator; Nuclear protein; Transcription regulation;
KW Repressor.
FT DOMAIN 11 69 CHROMO.
FT DOMAIN 383 395 POLY-HIS.
FT SEQUENCE 551 AA; 60581 MW; 30CEB09A82C58400 CRC64;
SQ
Query Match 14.9%; Score 76; DB 1; Length 551;
Best Local Similarity 37.8%; Pred. No. 1.1;
Matches 17; Conservative 2; Mismatches 18; Indels 8; Gaps 2;
QY 45 SRAGANHHHHHHGGH---GFVVRETRVEEDINTCTGEVHERRE 86
DB 378 SHPPAHHHHHHHHHHHTVGLNLSHAR-----RCLSETHGERE 417
RESULT 12
POUL_BRARE
ID POUL_BRARE STANDARD; PRT; 425 AA.
AC P31366;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POU domain protein 1 (ZFP001).
GN POU1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP MEDLINE-93038620; PubMed-1417821;
RX Matsuzaki T., Amanuma H., Takeda H.;
RT "A POU-domain gene of zebrafish, ZFP001, specifically expressed in the developing neural tissues."
RL Biochem. Biophys. Res. Commun. 187:1446-1453(1992).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN THE DEVELOPING NEURAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; AF055377; AAC27038.1; --
DR HSP; P23197; LAP0
DR MGD; MGI:1195985; Cbx4.
DR InterPro; IPR000953; Chromo.
DR Pfam; PF00385; chromo; 1.
DR PRINTS; PR00504; CHROMODOMAIN.
DR SMART; SM00298; CHROMO.1.
DR PROSITE; PS00598; CHROMO.1; 1.
DR PROSITE; PS00113; CHROMO.2; 1.
DR Chromatin regulator; Nuclear protein; Transcription regulation;
KW Repressor.
FT DOMAIN 11 69 CHROMO.
FT DOMAIN 383 395 POLY-HIS.
FT SEQUENCE 551 AA; 60581 MW; 30CEB09A82C58400 CRC64;
SQ
Query Match 14.9%; Score 76; DB 1; Length 403;
Best Local Similarity 31.6%; Pred. No. 0.8;
Matches 18; Conservative 8; Mismatches 25; Indels 6; Gaps 1;
QY 11 SEEVSVAPAGRHGGVQOH-----VVKFEFVDVTSRAGANHHHHHHGGH 61
DB 135 AQLAAAGAGAGASLGSGEEMGPAAVVSATIAAAGSAGPHHHHHHAAGH 191
RESULT 11
CBX4_MOUSE
ID CBX4_MOUSE STANDARD; PRT; 551 AA.
AC O55187;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromobox protein homolog 4 (Polycomb 2 homolog) (Pc2) (MPC2).
GN CBX4 OR PC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE-98035734; PubMed-9367786;
RX Alkema M.J., Jacobs J., Voncken J.W., Jenkins N.A., Copeland N.G., Satijn D.P.E., Otte A.P., Beris A., van Lohuizen M.;
RT "MPC2, a new murine homolog of the Drosophila polycomb protein is a member of the mouse polycomb transcriptional repressor complex."
RL J. Mol. Biol. 273:993-1003(1997).
CC -1- FUNCTION: INVOLVED IN MAINTAINING THE TRANSCRIPTIONALLY REPRESSIVE STATE OF GENES. MODIFIES CHROMATIN, RENDERING IT HERITABLY CHANGED IN ITS EXPRESSIBILITY.
CC -1- SUBUNIT: COMPONENT OF THE CHROMATIN-ASSOCIATED POLYCOMB COMPLEX (PCG).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 CHROMO DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
DR TRANSFAC; T01430; --
DR InterPro; IPR004826; TF_Maf.
DR InterPro; IPR004827; TF_bZip.
DR Pfam; PF03131; bZip_Maf; 1.
DR SMART; SM00338; BRU2; 1.
KW Oncogene; Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 163 167 POLY-ALA.
FT DOMAIN 173 182 HIS-RICH.
FT DOMAIN 184 191 POLY-GLY.
FT DOMAIN 214 223 POLY-GLY.
FT DOMAIN 233 240 POLY-GLY.
FT DOMAIN 274 300 BASIC MOTIF.
FT DNAME_BIND 302 323 LEUCINE-ZIPPER.
FT DNAME 369 AA; 38892 MW; F38GB220ACE50FF6 CRC64;
SQ SEQUENCE 369 AA; 38892 MW; F38GB220ACE50FF6 CRC64;

Query Match 14.7%; Score 75; DB 1; Length 369;
Best Local Similarity 37.8%; Pred. No. 0.93;
Matches 17; Conservative 2; Mismatches 18; Indels 8; Gaps 1;

QY 26 GGGVQVHVKEFEVDVTSRAGA-----NHHHHHHHGGHG 62
Db 143 GGSVPAEMGSAAVSAVIAAAAGGAPHYHHHHHHHPHGGGG 187

RESULT 15
CSUP DROME STANDARD; PRT; 449 AA.
AC Q9V3M4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catecholamines up protein.
GN Catsup OR CG10449.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99403013; PubMed=10471719;
RA Stathakis D.G., Burton D.Y., McIvor W.E., Krishnakumar S.,
RA Wright T.R., O'Donnell J.N.;
RA "The catecholamines up (Catsup) protein of Drosophila melanogaster
RA functions as a negative regulator of tyrosine hydroxylase activity.";
RA Genetics 153:361-362(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```

```
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC - FUNCTION: NEGATIVELY REGULATES TYROSINE HYDROXYLASE ACTIVITY.
CC - SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC - SIMILARITY: BELONGS TO THE KEA/CATSUP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF216584; AAF37226.1; --
CC EMBL; AE003661; AAF53744.1; --
CC FlyBase; FBgn0002022; Catsup.
CC InterPro; IPR002395; Kininogen.
CC InterPro; IPR003689; Zn_trnprt_zip.
CC Pfam; PF02535; Zip; 1.
CC PRINTS; PR00334; KININOGEN.
KW Transmembrane.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 167 187 POTENTIAL.
FT TRANSMEM 222 242 POTENTIAL.
FT TRANSMEM 371 391 POTENTIAL.
FT TRANSMEM 395 415 POTENTIAL.
FT CARBOHYD 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 449 AA; 48658 MW; F7111A254C07AB4C CRC64;

Query Match 14.5%; Score 74; DB 1; Length 449;
Best Local Similarity 32.7%; Pred. No. 1.4;
Matches 18; Conservative 7; Mismatches 20; Indels 10; Gaps 2;

QY 48 GANH-HHHGH-----HGCGFVFRVTRVEDINTCTGEVHERRESFLARA 92
Db 90 GHDGHHHHGHDHDDHGHGHHGHDHRTAKPDLDMSTIWLHSGSTLLISA 144

Search completed: June 6, 2003, 12:50:40
Job time : 14 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:47:53 ; Search time 50 Seconds
(without alignments)
383.248 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYOEVDCSEVRVAP.....INTCTGVEHRESFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	274	53.7	229	10 Q8S0B1	Q8S0B1 oryza sativ
2	177	34.7	95	10 Q9SM40	Q9SM40 sporobolus
3	141	27.6	102	10 Q8S0B2	Q8S0B2 oryza sativ
4	92.5	18.1	196	10 Q93ZA6	Q93ZA6 arabidopsis
5	92.5	18.1	196	10 Q9SM63	Q9SM63 arabidopsis
6	92.5	18.1	259	10 Q24643	Q24643 arabidopsis
7	89	17.5	1245	3 Q9P543	Q9P543 neurospora
8	86.5	17.0	408	10 Q8S0B4	Q8S0B4 oryza sativ
9	86.5	17.0	471	5 Q9UAY0	Q9UAY0 caenorhabd
10	86	16.9	793	4 Q9Y2L9	Q9Y2L9 homo sapien
11	84	16.5	441	10 Q944Q7	Q944Q7 arabidopsis
12	84	16.5	473	10 Q9FZ33	Q9FZ33 arabidopsis
13	84	16.5	1561	5 Q9W3D2	Q9W3D2 drosophila
14	83	16.3	168	5 Q9VWM5	Q9VWM5 drosophila
15	82.5	16.2	381	13 Q90WV0	Q90WV0 petromyzon
16	82.5	16.2	735	5 Q9NES7	Q9NES7 caenorhabd1

17	82	15.1	173	10 Q9LPS9	Q9LPS9 arabidopsis
18	81	15.9	86	5 Q20691	Q20691 caenorhabd1
19	81	15.9	109	11 Q9D6B9	Q9D6B9 mus musculus
20	81	15.9	191	10 Q42448	Q42448 medicago sa
21	81	15.9	397	5 Q9NC84	Q9NC84 strongyloce
22	80.5	15.8	347	5 Q9VX61	Q9VX61 drosophila
23	80.5	15.8	396	10 Q9FH24	Q9FH24 arabidopsis
24	80	15.7	133	10 Q9ZRI2	Q9ZRI2 medicago sa
25	80	15.7	1273	4 Q8Q316	Q8Q316 homo sapien
26	79	15.5	612	10 Q8RLS9	Q8RLS9 oryza sativ
27	79	15.5	612	10 Q8RUM5	Q8RUM5 oryza sativ
28	78	15.3	156	10 Q39754	Q39754 fagus sylv
29	78	15.3	585	5 Q9U4F0	Q9U4F0 drosophila
30	78	15.3	586	5 Q8T8P8	Q8T8P8 drosophila
31	78	15.3	990	5 Q9U4F1	Q9U4F1 drosophila
32	78	15.3	1020	5 Q9U1K1	Q9U1K1 drosophila
33	77.5	15.2	489	17 Q8THX7	Q8THX7 methanosarc
34	77.5	15.2	537	5 Q9VTJ6	Q9VTJ6 drosophila
35	77.5	15.2	566	5 Q9VTJ7	Q9VTJ7 drosophila
36	77.5	15.2	709	3 Q96W52	Q96W52 emeritella
37	77.5	15.2	1693	5 Q9VXV8	Q9VXV8 drosophila
38	77	15.1	354	5 Q8SSW8	Q8SSW8 dictyosteli
39	77	15.1	509	5 Q94888	Q94888 drosophila
40	77	15.1	509	5 Q9VZN8	Q9VZN8 drosophila
41	77	15.1	522	5 Q9XTK9	Q9XTK9 drosophila
42	77	15.1	834	6 Q9GKQ7	Q9GKQ7 ovis aries
43	76.5	15.0	77	5 Q20690	Q20690 caenorhabd1
44	76.5	15.0	332	16 Q9KG54	Q9KG54 bacillus ha
45	76.5	15.0	1283	5 Q95RH4	Q95RH4 drosophila

ALIGNMENTS

RESULT 1

Q8S0B1	PRELIMINARY;	PRT;	229 AA.
ID	Q8S0B1		
AC	Q8S0B1		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DE	P0470A12.5 protein.		
GN	P0470A12.5		
OS	Oryza sativa (japonica cultivar-group).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzaceae; Oryza.		
OX	NCBI_TaxID=39947;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC		
RT	clone:P0470A12.5";		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AP003436; BAB90280.1;		
SQ	SEQUENCE 229 AA; 25296 MW; D6EDAA65FFEF61E CRC64;		

Query Match 53.7%; Score 274; DB 10; Length 229;
Best Local Similarity 58.1%; Pred. No. 4.9e-23;
Matches 54; Conservative 14; Mismatches 21; Indels 4; Gaps 2;

QY	2	AYOEVDCSEVRVAP-AGFGRHGGVQVQVVKFEFVDVVSAGANHHHHHHGG 60	
DB	140	AQEVEVDCSEVRVTPGTGLGRGQVQVVKETFBQID---RSGSGRRHHHHNHN 196	
QY	61	HGFVVRVREEDINTCTGVEHRESFLARAN 93	
DB	197	DYLMVRETKVEEDNTCTGFEFRERKQSFLLKSD 229	

RESULT 2
Q9SM40

```

ID Q9SM40 PRELIMINARY; PRT; 95 AA.
AC Q9SM40;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Putative glycine-rich protein.
OS Sporobolus stapfianus (Ressurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Eragrostideae; Sporobolus.
NCBI_TaxID=56623;
RN [1]
SEQUENCE FROM N.A.
RA Neale A.D., Blomstedt C.K., Bronson P., Le T.N., Guthridge K.,
RA Evans J., Gaff D.F., Hamill J.D.;
RA "The isolation of lowly-transcribed genes which are induced during
RT desiccation of the resurrection grass Sporobolus stapfianus";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ242802; CAB61838.1; D756DCE2B68DD85B CRC64;
SQ SEQUENCE 95 AA; 10668 MW; 10668 MW; D756DCE2B68DD85B CRC64;

Query Match 34.7%; Score 177; DB 10; Length 95;
Best Local Similarity 57.7%; Pred. No. 1.5e-12;
Matches 45; Conservative 8; Mismatches 15; Indels 10; Gaps 6;

QY 1 MAYTQEVDCSEVRVAPAGFGRH-GGGVQOHVVKKEFEVDTVSRAGANHHHHHHH 58
DB 1 MAHFKEVMDY--EVTSMARKPFGRRGGGGVQGVVVKKEFEVEQVTPGRS-GHGGHR 57

QY 59 G--GHG---FVVRTRVE 71
DB 58 GNQGGSGHFOARETFE 75

RESULT 3
Q8SOB2 PRELIMINARY; PRT; 102 AA.
AC Q8SOB2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P0470A12.4 protein.
GN P0470A12.4.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0470A12." to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003436; BAB90279.1; F6F1266B1CDE7768 CRC64;
SQ SEQUENCE 102 AA; 11214 MW; 6FF1266B1CDE7768 CRC64;

Query Match 27.6%; Score 141; DB 10; Length 102;
Best Local Similarity 45.8%; Pred. No. 2e-08;
Matches 27; Conservative 8; Mismatches 18; Indels 6; Gaps 1;

QY 3 TYQEVDCSEVRVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHHGGH 61
DB 4 TYSEVDHCAEEMRPP-----HAGGEHYAVRRESTEVEDEMARAGRHHHNGGGGGH 56

RESULT 4
Q93ZA6 PRELIMINARY; PRT; 196 AA.
AC Q93ZA6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

ID Q9SM40 PRELIMINARY; PRT; 95 AA.
AC Q9SM40;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Putative glycine-rich protein.
OS Sporobolus stapfianus (Ressurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Eragrostideae; Sporobolus.
NCBI_TaxID=56623;
RN [1]
SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koeseema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057683; AAL15314.1;
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR ProDom; PD001423; TF_AP2; 1.
SQ SEQUENCE 196 AA; 21392 MW; E423D5570745C9CD CRC64;

Query Match 18.1%; Score 92.5; DB 10; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.013; 24; Indels 11; Gaps 5;
Matches 24; Conservative 12; Mismatches 12; Indels 11; Gaps 5;

QY 21 GFGRRGGG-VQOHVVKKEFE---VDTVSRAGA---NHHHHHHH-HGCHGVVRET--R 69
DB 99 GGGVNGGGDMSAAVIRRKAAEVGAQVDALEAGAGNRHHHHHHOHRGNDYVDNHSYR 158

QY 70 VEEDINTCTGE 80
DB 159 INDDLMECSK 169

RESULT 5
Q9SW63 PRELIMINARY; PRT; 196 AA.
AC Q9SW63;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE TINY-like protein.
GN C7A10.460 OR A74G36900.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Mayer K.F.X.,
RA Schueller C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99707; CAB16766.1;
DR EMBL; AL161590; CAB80356.1;
DR HSPF; O80337; ZGCC
DR InterPro; IPR001471; TF_ERF
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSEPELMNT.
DR ProDom; PD001423; TF_AP2; 1.
DR SMART; SM00380; AP2; 1.
```

```

DE  Related to multifunctional cyclin-dependent kinase PHO85.
GN  B24H17.10.
OS  Neurospora crassa.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC  Sordariales; Sordariaceae; Neurospora.
OX  NCBI_TaxID=5141;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA  Nyakatura G., Mewes H.W., Mannhaupt G.;
RL  Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  German Neurospora genome project;
RL  Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL  EMBL; AL356815; CAB92623.1; -
DR  HSSP; P42773; 1IHB.
DR  InterPro; IPR002110; ANK.
DR  InterPro; IPR004129; GPD.
DR  InterPro; IPR004331; SPX.
DR  Pfam; PF00023; ank; 6.
DR  Pfam; PF03049; GPDF; 1.
DR  Pfam; PF03105; SPX; 1.
DR  SMART; SM00248; ANK; 3.
DR  PROSITE; PS50088; ANK_REPEAT; 3.
DR  PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW  ANK repeat; Cyclin; Kinase; Repeat.
SQ  SEQUENCE 1245 AA; 137798 MW; 16BD0F6A04596A9F CRC64;

Query Match 17.5%; Score 89; DB 3; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.29;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  45 SRAGANHHHHHGGHGG 62
DB  901 SRSGTHHHHHHHGGH 918
    ||.||:||||| |||| ||
    ||:||||| |||| ||||

RESULT 8
Q850B4
ID  Q850B4 PRELIMINARY; PRT; 408 AA.
AC  Q850B4;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  P0470A12.2 protein.
DE  P0470A12.2.
OS  Oryza sativa (japonica cultivar-group).
OC  Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzeae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV. NIPPONBARE;
RA  Sasaki T., Matsumoto T., Yamamoto K.;
RT  "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT  clone:P0470A12.2."
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL  EMBL; AP003436; BAB90277.1; -
SQ  SEQUENCE 408 AA; 43466 MW; F7E0297877F6D69C CRC64;

Query Match 17.0%; Score 86.5; DB 10; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.15;
Matches 25; Conservative 10; Mismatches 15; Indels 43; Gaps 5;

QY  12 EEVRSVAPAGFGRH-----GGGQQHVVKFEVDVTSRAGA-----NH----- 52
DB  199 EECDAAGGGYGRHHCAGGAVKQHATYKQHQAIQGVNGCCAGGYNRHOAVAGGG 258
    ||:||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
    ||:||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY  53 HH-----GH--GGH 61
DB  259 HHYGGATAAAYGNASNKQHFATAAGHHSSGGH 291
    ||:||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

```

SEQUENCE FROM N.A.
 TISSUE-BRAIN;
 RC MEDLINE=99246063; PubMed=10231032; Kikuno R., Hirose M.,
 RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Ohara O.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura O.,
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro."
 RL DNA Res. 6:63-70(1999).
 DR EMBL; AB023233; BAA76860.1; -
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002114; HPR_SerP_site.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00307; CH; 1.
 DR Pfam; PF00560; LRR; 6.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00370; LRR; 3.
 DR SMART; SM00369; LRR_TYP; 1.
 DR PROSITE; PS00021; CH; 1.
 DR PROSITE; PS00589; PHS_HPR_SER; UNKNOWN_1.
 DR NON_TER
 SQ SEQUENCE 793 AA; 87801 MW; 0BE99CC48C3BB37C CRC64;
 Query Match 16.9%; Score 86; DB 4; Length 793;
 Best Local Similarity 32.8%; Pred. No. 0.37;
 Matches 22; Conservative 7; Mismatches 28; Indels 10; Gaps 2;
 QY 6 EVDYCESEVRSVAPAGFRHGGVQOHVYKFEF---VDTVSRAGA-----NHHHHH 55
 DB 39 ELFPSPRGECDDPPAGAGRGGAKEKMATPGEPPVPSALSVATLPLHHPHHHHH 98
 QY 56 GHGGHG 62
 DB 99 QHGGTG 105
 RESULT 11
 Q944Q7 PRELIMINARY; PRT; 441 AA.
 ID Q944Q7
 AC Q944Q7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE At1g54990/F14C21.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF424608; AAL11602.1;
 SQ SEQUENCE 441 AA; 48816 MW; 520163FE0A8DE447 CRC64;
 Query Match 16.5%; Score 84; DB 10; Length 441;
 Best Local Similarity 36.8%; Pred. No. 0.31;
 Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;
 QY 8 DYCESEVRSVAPAGFRHGGVQOHVYKFEF---VDTVSRAGANHHHHH---HGHH 61
 DB 39 ELFPSPRGECDDPPAGAGRGGAKEKMATPGEPPVPSALSVATLPLHHPHHHHH 98

SEQUENCE FROM N.A.
 Q9UAY0 PRELIMINARY; PRT; 471 AA.
 ID Q9UAY0
 AC Q9UAY0
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE W03G1.5 protein.
 GN W03G1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Pauley A., Scheet P., Harper M.;
 RT "The sequence of C. elegans cosmid W03G1.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Waterston R.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF125964; AAD14753.1;
 SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A985B CRC64;
 Query Match 17.0%; Score 86.5; DB 5; Length 471;
 Best Local Similarity 34.8%; Pred. No. 0.17; 16; Indels 25; Gaps 4;
 Matches 24; Conservative 4; Mismatches 4;
 QY 15 RSVAPAGFRHGG-----GVQOHVYKFEFVDTVSRAGANHHHHH--- 56
 DB 372 RSHSPRGHG-HGGRGPHPCFRGHGPHPH-----HHHDGRSPSRHGHHHHHHGGCRPF 426
 QY 57 --HHGGHGF 63
 DB 427 PPHGHGHHF 435
 RESULT 10
 Q9Y2L9 PRELIMINARY; PRT; 793 AA.
 ID Q9Y2L9
 AC Q9Y2L9
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE KIAA1016 protein (Fragment).
 GN KIAA1016.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]

db 377 DYISEFV-SLLPKSIRR----VAEPIPEEVQKVL EEAKAGDDHDDHHGHGHAGY 428

RESULT 12

Q9FZ33 PRELIMINARY; PRT; 473 AA.

01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
T24C10.10 protein (Hypothetical 52.4 kDa protein).
T24C10.10 OR F14C21.51 OR A1IG54990.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altati H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chiu J., Choi E., Gonzalez A.,
RA Howing B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsky N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Ecker J.R., Theologis A., Davis R.W.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.B., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Matti R., Marzilli A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT *Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana.;
RA Nature 408:816-820(2000).
[3]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT *Full length cDNA of gene Atlg54990 (GI:15221965).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RA EMBL; AC064840; AAG00878.1; -
RA EMBL; AC069144; AAG51115.1; -
RA EMBL; AY074334; AAL67030.1; -
RA Hypothetical protein.
SEQUENCE 473 AA; 52419 MW; D09124A11565BB23 CRC64;

```
Query Match          16.5%; Score 84; DB 10; Length 473;  
Best Local Similarity    36.8%; Pred. No. 0.34;  
Matches      21; Conservative   10; Mismatches     18; Indels       8; Gaps        3;
```

2Y

```
      8 DYCSEEVRSVAPAGFCRHGGGVQOHHVVVKFEVDTVSRAGANHHHHHGH---RGHG 61  
|| ||| | : | : | : | : | : | : | : | : | : | : |  
|| ||| | : | : | : | : | : | : | : | : | : | : |
```

Db 409 DYISEFV-SLLPKSIRR---VAEPIPEEVQKVL EAKAGDDHHHGHGHAGY 460

RESULT 13

ID	Q9W3D2	PRELIMINARY;	PRT; 1561 AA.
AC	Q9W3D2;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	CRAG protein.		
DE	CRAG OR CG12737.		
GN	Drosophila melanogaster (Fruit fly).		
OS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Ephytrida; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephytridae; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,		
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeoon K.Y., Benos P.V., Berzman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotter P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moyle M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,		
RA	Palazolo M., Pittman G.S., Pan S., Pollard J.C., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster";		
RL	Science 287:2185-2195(2000).		
DR	EMBL; AE003444; AAF46397.1; -.		
DR	FlyBase; FBgn00258664; Crag.		
DR	InterPro; IPR005112; dDENN.		
DR	InterPro; IPR001194; DENN.		
DR	InterPro; IPR005113; dDENN.		
DR	Pfam; PF03455; dDENN; 1.		
DR	Pfam; PF02141; DENN; 1.		
DR	Pfam; PF03456; uDENN; 1.		
Q9W3D2	SEQUENCE 1561 AA; 174616 MW; 10BFDD38A3DD4FC4E CRC64;		

Query Match 16.5%; Score 84; DB 5; Length 1561;
Best Local Similarity 23.9%; Pred. NO. 1.4;
Matches 28; Conservative 17; Mismatches 30; Indels 42; Gaps 5;

Search completed: June 6, 2003, 12:52:27
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:02:07 ; Search time 1293.5 seconds
(without alignments)
1164.422 Million cell updates/sec

Title: US-10-090-035-2

Perfect score: 510

Sequence: 1 MAYQVDYGVSEVRVAPV.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUN62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+...p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10090035/runat_06062003_105504_10970/app.query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10090035@cgn_1_1_2463@runat_06062003_105504_10970 -NCPU=6 -ICPU=3
-NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_hic:.*
9: gb_estl:.*
10: gb_est2:.*
11: gb_hic:.*
12: gb_est3:.*
13: gb_est4:.*
14: gb_est5:.*
15: em_estfun:.*
16: em_estom:.*
17: gb_gss:.*
18: em_gss_hum:.*
19: em_gss_inv:.*
20: em_gss_pln:.*
21: em_gss_vrt:.*
22: em_gss_fun:.*
23: em_gss_mam:.*
24: em_gss_mus:.*
25: em_gss_other:.*
26: em_gss_pro:.*
27: em_gss_rod:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	417	13	BM501439
2	505	99.0	465	10	BE129897
3	505	99.0	500	10	AW787732
4	505	99.0	523	10	AW331212
5	505	99.0	524	10	AW288876
6	499.5	97.9	466	9	AI964534
7	499.5	97.9	481	9	AI964458
8	499.5	97.9	539	9	AI855425
9	499.5	97.9	546	9	AA979839
10	493.5	97.9	550	12	BG840383
11	499.5	97.9	553	11	AY104409
12	499.5	97.9	648	14	BQ619167
13	499.5	97.9	648	14	BQ619315
14	499.5	97.9	648	14	BQ619318
15	499.5	97.9	648	14	BQ619337
16	499.5	97.9	648	14	BQ619338
17	499.5	97.9	648	14	BQ619390
18	498	97.6	326	10	BE025303
19	496	97.3	443	10	BE129644
20	496	97.3	552	10	BE025302
21	494	96.9	476	10	BE519299
22	492	96.5	496	10	AW787314
23	482	94.5	433	10	BE225008
24	468	91.8	436	10	AW787315
25	423.5	83.0	399	13	BM318672
26	423.5	83.0	466	10	BE597738
27	423.5	83.0	485	10	BE599123
28	423.5	83.0	513	14	BQ280709
29	423.5	83.0	516	10	BE364814
30	423.5	83.0	519	10	AW680016
31	423.5	83.0	522	10	AW679969
32	423.5	83.0	537	10	AW745400
33	423.5	83.0	541	10	AW745436
34	417.5	81.9	348	10	BE593507
35	410.5	80.5	509	14	BQ280894
36	395	77.5	401	10	BE289056
37	368.5	72.3	274	10	AW288875
38	367.5	72.1	311	12	BF729420
39	308	60.4	453	10	AW923922
40	308	60.4	554	10	AW679915
41	308	60.4	560	10	AW924079
42	308	60.4	566	10	AW677917
43	308	60.4	566	10	AW746383
44	308	60.4	572	10	AW679713
45	308	60.4	572	10	AW925014

ALIGNMENTS

RESULT 1
BM501439
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 417)
Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and Jung,R.
BM501439
PAC000000000593 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
417 bp mRNA linear EST 14-FEB-2002
BM501439
EST.
Zea mays
Zea mays
BM501439.1 GI:18661517
EST.
Zea mays

Maize opaque endosperm mutations create extensive changes in patterns of gene expression
Unpublished (2002)
Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.

FEATURES
source
1. .465
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site 1: ECORI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 88 a 151 c 149 g . 77 t
ORIGIN

Alignment Scores:
Pred. No.: 2,26e-46 Length: 465
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-2 (1-93) x BE129897 (1-465)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 40 ATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTGGTGGCCCGGCC 99
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 100 GGCCTTCGGCGCCACCGCGCGCTCCAGCAGCAGCTCGTCAGGAGAGTTCGAGGAG 159
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisGlyHisGlyGly 60
Db 160 GTCGACACGGTCTACGCGCGCGCGCCACACACACACCATGTCACCGCGCGC 219
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
Db 220 CACGGCTTCGTGGTGGCGGAGCAGGAGTCCAGGAGGACATCAACACCTGCACCGCGCAG 279
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 280 TTCACGAGCGCAGGAGAGAGCTTCCTCCGCGAGGCTAAC 318

RESULT 3
AW787732 500 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787732
VERSION AW787732.1 GI:7844510
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 500)
REFERENCE Walbot,V.
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Maize opaque endosperm mutations create extensive changes in patterns of gene expression
Unpublished (2002)
Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.

FEATURES
source
1. .417
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/notes="Vector: pSport1; Site 1: SalI; Site 2: NotI"

BASE COUNT 81 a 150 c 133 g 53 t
ORIGIN

Alignment Scores:
Pred. No.: 5,6e-47 Length: 417
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-090-035-2 (1-93) x BM501439 (1-417)

QY 1 MetAlaTyrrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 85 ATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAGTGGTGGCCCGGCC 144
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 145 GGCCTTCGGCGCCACCGCGCGCTCCAGCAGCAGCTCGTCAGGAGAGTTCGAGGAG 204
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisGlyHisGlyGly 60
Db 205 GTCGACACGGTCTACGCGCGCGCGCCACACACCATGTCACCGCGCGC 264
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
Db 265 CACGGCTTCGTGGTGGCGGAGCAGGAGTCCAGGAGGACATCAACACCTGCACCGCGCAG 324
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 325 GTCACGAGCGCAGGAGAGAGCTTCCTCCGCGAGGCTAAC 363

RESULT 2
BE129897 465 bp mRNA linear EST 21-JUN-2000
LOCUS 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897
VERSION BE129897.1 GI:8577260
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 465)
REFERENCE Walbot,V.
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.

FEATURES
SOURCE

```

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mix
same as 707 (SK)"
/tissue_type="tassel,
/dev_stage="fully-gro
/lab host="PH108"

```

/note="Organ:tassel, kernal, silk, husk, root, leaf;
vector:pGD10; Site+: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:3:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

**BASE COUNT
ORIGIN**

given to library 707 for ad
a 159 c 156 g 88 t

Alignment Scores:

Pred. No.:	2,458-46	Length:	500
Score:	505.00	Matches:	92
Percent Similarity:	98.9%	Conservative:	0
Best Local Similarity:	98.9%	Mismatches:	1
Query Match:	99.0%	Indels:	0
DB:	10	Gaps:	0

US-10-090-035-2 (1-93) x AW787732 (1-500)

1	Meta	Ala	Tyr	Tyr	Gln	Glu	Val	Asp	Tyr	Cys	Ser	Glu	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20
35	ATG	CTT	ACT	ACC	AGG	AGT	GACT	ACT	TGCT	CGG	AGG	AGT	GAGG	TGCG	GGT	GGCC	CGGCC	94			
21	Gly	Phe	Gly	Arg	His	Gly	Gly	Val	Gln	Gln	His	Val	Val	Lys	Glu	Lys	Phe	Glu	Glu	40	
95	GGC	TTG	GGC	CCG	CAC	GGG	GGG	GTC	CAC	AGC	AGC	CTG	CTG	AAG	GAG	AAG	TTT	CGA	GAG	154	
41	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	His	His	His	His	Gly	His	His	Gly	Gly	60	
155	GTC	GAC	CAG	GGT	GCT	CAC	GGC	GGG	CGC	CA	CAC	CAC	CAC	CAC	CAC	CAT	TGT	CAC	CAC	214	
61	His	Gly	Phe	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Glu	Asp	Leu	Asn	Thr	Cys	Thr	Gly	80	
215	CAC	GGC	TCG	TGT	GGC	GAG	CAC	GGG	TG	CA	AGG	AGC	ATC	AAC	ACC	TGC	ACC	CGG	CAG	274	
81	Val	His	Glu	Arg	Arg	Glu	Ser	Phe	Leu	Ala	Arg	Ala	Asn	93							
275	TT	CC	GAG	CGC	AGG	GAG	AGT	TC	CT	CGC	AGG	CT	TAAC	313							

RESULT 4
AW331212

LOCUS	523 bp	mRNA	linear	EST 31-JAN-2000
DEFINITION	707049E04.x1 707 - Mixed adult tissues from Walbot lab (SK) Zea mays cDNA, mRNA sequence.			

ACCESSION AW331212
VERSION AW331212.1
KEYWORDS EST.

SOURCE	Zea mays.
ORGANISM	Zea mays
1. <i>Aspergillus fumigatus</i>	1. <i>Aspergillus fumigatus</i>
2. <i>Aspergillus niger</i>	2. <i>Aspergillus niger</i>
3. <i>Aspergillus terreus</i>	3. <i>Aspergillus terreus</i>
4. <i>Aspergillus oryzae</i>	4. <i>Aspergillus oryzae</i>
5. <i>Aspergillus nidulans</i>	5. <i>Aspergillus nidulans</i>
6. <i>Aspergillus glaucus</i>	6. <i>Aspergillus glaucus</i>
7. <i>Aspergillus clavatus</i>	7. <i>Aspergillus clavatus</i>
8. <i>Aspergillus carbonum</i>	8. <i>Aspergillus carbonum</i>
9. <i>Aspergillus versicolor</i>	9. <i>Aspergillus versicolor</i>
10. <i>Aspergillus nidulans</i>	10. <i>Aspergillus nidulans</i>
11. <i>Aspergillus fumigatus</i>	11. <i>Aspergillus fumigatus</i>
12. <i>Aspergillus niger</i>	12. <i>Aspergillus niger</i>
13. <i>Aspergillus terreus</i>	13. <i>Aspergillus terreus</i>
14. <i>Aspergillus oryzae</i>	14. <i>Aspergillus oryzae</i>
15. <i>Aspergillus nidulans</i>	15. <i>Aspergillus nidulans</i>
16. <i>Aspergillus glaucus</i>	16. <i>Aspergillus glaucus</i>
17. <i>Aspergillus clavatus</i>	17. <i>Aspergillus clavatus</i>
18. <i>Aspergillus carbonum</i>	18. <i>Aspergillus carbonum</i>
19. <i>Aspergillus versicolor</i>	19. <i>Aspergillus versicolor</i>
20. <i>Aspergillus nidulans</i>	20. <i>Aspergillus nidulans</i>
21. <i>Aspergillus fumigatus</i>	21. <i>Aspergillus fumigatus</i>
22. <i>Aspergillus niger</i>	22. <i>Aspergillus niger</i>
23. <i>Aspergillus terreus</i>	23. <i>Aspergillus terreus</i>
24. <i>Aspergillus oryzae</i>	24. <i>Aspergillus oryzae</i>
25. <i>Aspergillus nidulans</i>	25. <i>Aspergillus nidulans</i>
26. <i>Aspergillus glaucus</i>	26. <i>Aspergillus glaucus</i>
27. <i>Aspergillus clavatus</i>	27. <i>Aspergillus clavatus</i>
28. <i>Aspergillus carbonum</i>	28. <i>Aspergillus carbonum</i>
29. <i>Aspergillus versicolor</i>	29. <i>Aspergillus versicolor</i>
30. <i>Aspergillus nidulans</i>	30. <i>Aspergillus nidulans</i>
31. <i>Aspergillus fumigatus</i>	31. <i>Aspergillus fumigatus</i>
32. <i>Aspergillus niger</i>	32. <i>Aspergillus niger</i>
33. <i>Aspergillus terreus</i>	33. <i>Aspergillus terreus</i>
34. <i>Aspergillus oryzae</i>	34. <i>Aspergillus oryzae</i>
35. <i>Aspergillus nidulans</i>	35. <i>Aspergillus nidulans</i>
36. <i>Aspergillus glaucus</i>	36. <i>Aspergillus glaucus</i>
37. <i>Aspergillus clavatus</i>	37. <i>Aspergillus clavatus</i>
38. <i>Aspergillus carbonum</i>	38. <i>Aspergillus carbonum</i>
39. <i>Aspergillus versicolor</i>	39. <i>Aspergillus versicolor</i>
40. <i>Aspergillus nidulans</i>	40. <i>Aspergillus nidulans</i>
41. <i>Aspergillus fumigatus</i>	41. <i>Aspergillus fumigatus</i>
42. <i>Aspergillus niger</i>	42. <i>Aspergillus niger</i>
43. <i>Aspergillus terreus</i>	43. <i>Aspergillus terreus</i>
44. <i>Aspergillus oryzae</i>	44. <i>Aspergillus oryzae</i>
45. <i>Aspergillus nidulans</i>	45. <i>Aspergillus nidulans</i>
46. <i>Aspergillus glaucus</i>	46. <i>Aspergillus glaucus</i>
47. <i>Aspergillus clavatus</i>	47. <i>Aspergillus clavatus</i>
48. <i>Aspergillus carbonum</i>	48. <i>Aspergillus carbonum</i>
49. <i>Aspergillus versicolor</i>	49. <i>Aspergillus versicolor</i>
50. <i>Aspergillus nidulans</i>	50. <i>Aspergillus nidulans</i>
51. <i>Aspergillus fumigatus</i>	51. <i>Aspergillus fumigatus</i>
52. <i>Aspergillus niger</i>	52. <i>Aspergillus niger</i>
53. <i>Aspergillus terreus</i>	53. <i>Aspergillus terreus</i>
54. <i>Aspergillus oryzae</i>	54. <i>Aspergillus oryzae</i>
55. <i>Aspergillus nidulans</i>	55. <i>Aspergillus nidulans</i>
56. <i>Aspergillus glaucus</i>	56. <i>Aspergillus glaucus</i>
57. <i>Aspergillus clavatus</i>	57. <i>Aspergillus clavatus</i>
58. <i>Aspergillus carbonum</i>	58. <i>Aspergillus carbonum</i>
59. <i>Aspergillus versicolor</i>	59. <i>Aspergillus versicolor</i>
60. <i>Aspergillus nidulans</i>	60. <i>Aspergillus nidulans</i>
61. <i>Aspergillus fumigatus</i>	61. <i>Aspergillus fumigatus</i>
62. <i>Aspergillus niger</i>	62. <i>Aspergillus niger</i>
63. <i>Aspergillus terreus</i>	63. <i>Aspergillus terreus</i>
64. <i>Aspergillus oryzae</i>	64. <i>Aspergillus oryzae</i>
65. <i>Aspergillus nidulans</i>	65. <i>Aspergillus nidulans</i>
66. <i>Aspergillus glaucus</i>	66. <i>Aspergillus glaucus</i>
67. <i>Aspergillus clavatus</i>	67. <i>Aspergillus clavatus</i>
68. <i>Aspergillus carbonum</i>	68. <i>Aspergillus carbonum</i>
69. <i>Aspergillus versicolor</i>	69. <i>Aspergillus versicolor</i>
70. <i>Aspergillus nidulans</i>	70. <i>Aspergillus nidulans</i>
71. <i>Aspergillus fumigatus</i>	71. <i>Aspergillus fumigatus</i>
72. <i>Aspergillus niger</i>	72. <i>Aspergillus niger</i>
73. <i>Aspergillus terreus</i>	73. <i>Aspergillus terreus</i>
74. <i>Aspergillus oryzae</i>	74. <i>Aspergillus oryzae</i>
75. <i>Aspergillus nidulans</i>	75. <i>Aspergillus nidulans</i>
76. <i>Aspergillus glaucus</i>	76. <i>Aspergillus glaucus</i>
77. <i>Aspergillus clavatus</i>	77. <i>Aspergillus clavatus</i>
78. <i>Aspergillus carbonum</i>	78. <i>Aspergillus carbonum</i>
79. <i>Aspergillus versicolor</i>	79. <i>Aspergillus versicolor</i>
80. <i>Aspergillus nidulans</i>	80. <i>Aspergillus nidulans</i>
81. <i>Aspergillus fumigatus</i>	81. <i>Aspergillus fumigatus</i>
82. <i>Aspergillus niger</i>	82. <i>Aspergillus niger</i>
83. <i>Aspergillus terreus</i>	83. <i>Aspergillus terreus</i>
84. <i>Aspergillus oryzae</i>	84. <i>Aspergillus oryzae</i>

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 523)

AUTHORS	TITLE
Walbot, V.	Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL
Unpublished (1999)

COMMENT
Contact: Walbot v
Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Place: 707049 row: E column: 04.

FEATURES

```

1. 523
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
) )
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pAD10; Site:1. Ecoli: cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

```

BASE COUNT
ORIGIN

mask, tool, real):	undire
a 158 c 149 g	91 t

Alignment Scores:

Pred. No.:	2,588-46	Length:	523
Score:	505.00	Matches:	92
Percent Similarity:	98.92%	Conservative:	0
Best Local Similarity:	98.92%	Mismatches:	1
Query Match:	99.02%	Indels:	0
DB:	10	Gaps:	0

US-10-090-035-2 (1-93) x AW331212 (1-523)

1	QY	Meta	Ty	Tyr	Gln	Glu	Val	Asp	Tyr	Cys	Ser	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20
35	Db	ATGG	CTT	TACT	TAC	CAG	GAG	TGG	ACT	TACT	GCT	CG	GAG	GAG	TGG	TGG	TGG	CGCC	GGCC	94
21	QY	Gly	Phe	Gly	Arg	His	Gly	Gly	Val	Gln	Gln	His	Val	Val	Lys	Glu	Lys	Phe	Glu	40
95	Db	GGCT	TCGG	CGCC	CAC	GGCG	GGCG	GCT	CAG	CAG	CAC	GCT	GCT	CA	GAG	GAG	AAG	TTC	GAG	155
41	QY	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	His	His	His	Gly	His	His	Gly	Gly	60
155	Db	GTC	GAC	ACG	GTCT	CAC	GGCG	GGCG	GCC	AAAC	CACC	CACC	ACC	ACC	TGT	CTAC	CAC	CGCG	CGC	215
61	QY	His	Gly	Phe	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Ala	Ser	Phe	Leu	Ala	Arg	Glu	80
215	Db	CAC	GGCT	TCGT	GTGT	CGCG	GAG	ACC	CAG	GGT	CG	AAG	GAC	ATC	CAAC	ACCT	CGAC	CGGG	CGAG	275
81	QY	Val	His	Glu	Arg	Arg	Glu	Ser	Phe	Leu	Ala	Arg	Ala	Asn						93
275	Db	TTC	CAC	GAG	CGC	GAG	GAG	AGT	TCT	CGC	CAC	AGG	GT	TAAC						313

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

RESULT 5

AW288876 LOCUS 524 bp mRNA linear EST 16-JAN-2000
DEFINITION 70709E07.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.

ACCESSION AW288876
VERSION AW288876.1
KEYWORDS EST GI:6695663

SOURCE
Zea mays.

ORGANISM *Zea mays*

Eukaryota
Spermatonspecimen of
clade; pa

REFERENCE	1 (bases)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AUTHORS	TITLE
Walbot, V.	Maize EST

University of Illinois at Chicago

JOURNAL
Unpublished

COMMENT:

Contact:

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496013 row: D column: 01.

FEATURES
source
1. .466
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E.Coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+)"
Wang/Bohnert"

BASE COUNT 76 a 151 c 154 g 85 t

ORIGIN

Alignment Scores:
Pred. No.: 9.2e-46 Length: 466
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI964534 (1-466)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 416 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGGAGGTGAGTCTGGTGGCCCGGCC 357

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 356 GGCCTTCGGCCGCCAGCGCGCGGCGTCCAGCAGCAGCGTCTCAAGAGAGTTCAGGAG 297

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisGlyHisGly 59
DB 296 GTCGACACGGCTCTACGCGCGCGGCCAACCCACCACCATGCTGTCACCCAGGC 237

QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
DB 236 GGCACGGCTTCGTGTGCGCGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 177

QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
DB 176 GAGTCCAGCGCGCAGGAGAGGTCTCTCGCCAGGCTAAC 135

RESULT 7
AI964458/c 481 bp mRNA linear EST 20-AUG-1999
LOCUS lab Zea mays cDNA, mRNA sequence.
DEFINITION 496010H04.x1 496 - stressed shoot cDNA library from Wang/Bohnert
ACCESSION AI964458
VERSION AI964458.1 GI:5757171
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 481)
REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES
source
1. .524
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH108"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 107 a 170 c 157 g 89 t

ORIGIN

Alignment Scores:
Pred. No.: 2.58e-46 Length: 524
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-2 (1-93) x AW288876 (1-524)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 73 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGGAGGTGAGTCTGGTGGCCCGGCC 132

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 133 GGCCTTCGGCCGCCAGCGCGCGGCGTCCAGCAGCAGCGTCTCAAGAGAGTTCAGGAG 192

QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisGlyHisGly 60
DB 193 GTCGACACGGCTCTACGCGCGCGGCCAACCCACCACCATGCTGTCACCCAGGC 252

QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 253 CACGGCTTCGTGTGCGCGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCGCGAG 312

QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 313 TTCCACAGCGCGCAGGAGAGGTCTCTCGCCAGGCTAAC 135

RESULT 6
AI964534/c 466 bp mRNA linear EST 20-AUG-1999
LOCUS lab Zea mays cDNA, mRNA sequence.
DEFINITION 496013D01.x1 496 - stressed shoot cDNA library from Wang/Bohnert
ACCESSION AI964534
VERSION AI964534.1 GI:5757247
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 466)
REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 723 2221
Email: walbot@stanford.edu
Plate: 496010 row: H column: 04.
Location/Qualifiers

FEATURES

source

1. 481
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+); XR; Wang/Bohnert"

BASE COUNT 82 a 148 c 154 g 97 t

ORIGIN

Alignment Scores:
Pred. No.: 9,53e-46 Length: 481
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI964458 (1-481)

QY 1 MetalatYrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 451 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGAGGAGTGGTGGTGGCCCGGCC 392
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 391 GCTTCGGCCGCCAGGAGGCGGCTCCAGCAGCAGTCTCGTCAAGAGGAAGTTCGAGGAG 332
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59
Db 331 GTCGACACGGTCTCACGGCGCGCGCCACACACACACACACACATGTCACACGGC 272
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 271 GCCCAGCGCTTCGTGGTGGCGAGACACAGGTCGAGGAGGACATCAACACCTGCACGGC 212
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 211 GAGGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 170

RESULT 8

AI855425/c
LOCUS 539 bp mRNA linear EST 15-JUL-1999
DEFINITION 603016F02.x1 603 - stressed root cDNA library from Wang/Bohnert lab
Zea mays cDNA, mRNA sequence.
ACCESSION AI855425
VERSION AI855425.1 GI:5499558
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 539)

REFERENCE
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.
Location/Qualifiers

FEATURES

source

1. 539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+); XR; Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 91 a 157 c 172 g 119 t

ORIGIN

Alignment Scores:
Pred. No.: 1.08e-45 Length: 539
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AI855425 (1-539)

QY 1 MetalatYrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 477 ATGGCTTACTACAGAGGTGGACTACTGCTCGAGAGGAGTGGTGGTGGCCCGGCC 418
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 417 GGTTCGGCGGCCAGGAGGCGGCTCCAGCAGCAGTCTCGTCAAGAGGAAGTTCGAGGAG 358
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisGly 59
Db 357 GTCGACACGGTCTCACGGCGCGCGCCACACACACACACACATGTCACACGGC 298
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 297 GCCCAGCGCTTCGTGGTGGCGAGACACAGGTCGAGGAGGACATCAACACCTGCACGGC 238
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 237 GAGGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 196

RESULT 9

AA979839
LOCUS 546 bp mRNA linear EST 26-MAY-1998
DEFINITION MEST2-B7.TW1412.Seq ISUM2 Zea mays cDNA clone MEST2-B7 5', mRNA sequence.
ACCESSION AA979839
VERSION AA979839.1 GI:3157217
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 546)

REFERENCE
AUTHORS Wen, T.J., Ashlock, D.A. and Schnable, P.S.
TITLE Expressed Sequence Tags from B73 Maize Seedlings Unpublished (1997)
JOURNAL Contact: Schnable, PS
COMMENT Schnable Laboratory

Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975
Fax: (515)-294-2299
Email: schnable@iastate.edu

Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206705.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

PCR Primers
FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T7-1 (AA TAC GAC TCA CTA TAG).

FEATURES
source
1..550
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST12-H11"
/clone_lib="ISUM4-TN"
/tissue_type="seedling and silk"
/lab_host="DH10B"
/note="Vector: pT7T3PAC; Site.1: EcoRI; Site.2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'
AACTGGAAGAATTCGCGCCGAGCAATTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT7T3PAC vector. The library
then went through one round of normalization to Cot value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

BASE COUNT 124 a 174 c 159 g 92 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1,11e-45 Length: 550
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 12 Gaps: 1

US-10-090-035-2 (1-93) x BG840383 (1-550)

QY 1 MetAlaTyrTyrGlnGluValAlaSerGluGluValAlaProAla 20
|||||
Db 75 ATGGCTTACTACAGAGGTTGGACTACTCTCGGAGGAGTGAGTGGTGGCCCGGCC 134
|||||

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 135 GGCTTCGGCGCCACGAGGCGCGTCCAGCAGCAGTCTCAAGGAGAGTTCGAGGAG 194
|||||

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
|||||
Db 195 GTCGACACGGTCTCAGCGCGCGCGCCACACCCACCCACCATGGTACACCGGC 254
|||||

QY 60 GlyHisGlyPheValValArgGluThrArgValGluAspIleAsnThrCysThrGly 79
|||||
Db 255 GGCCACGGCTTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 314
|||||

QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 315 GAGGTCCACGAGCGAGGAGAGTCTCTCCGCCAGGGCTAAC 356
|||||

RESULT 11
AY104409

PCR Primers
FORWARD: tw1412 (5'-GAAGATACCCCAACACC-3')
BACKWARD: T7-YJ (5'-TAATACGACTACTATAGGC-3')
Plate: MEST2 row: B column: 7
Seq primer: tw1412 (5'-CAAGATACCCCAACACC-3').

FEATURES
source
1..546
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST2-B7"
/clone_lib="ISUM2"
/tissue_type="above ground tissues"
/dev_stage="two-leaf-stage green seedling"
/lab_host="XLI-MFR Blue"
/note="Organ: green seedlings; Vector: pAD-GAL4; Site.1:
EcoRI; Site.2: XhoI; ds-cDNA molecules were generated as
follows. First-strand cDNA was prepared from oligo-dT
selected mRNA by priming with an XhoI oligo-dT primer. The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with XhoI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and XhoI sites of the HybriZAP lambda vector
(Stratagene) and excised as pAD-GAL4 phagemids."

BASE COUNT 130 a 168 c 159 g 88 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.1e-45 Length: 546
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-2 (1-93) x AA979839 (1-546)

QY 1 MetAlaTyrTyrGlnGluValAlaSerGluGluValAlaProAla 20
|||||
Db 66 ATGGCTTACTACAGAGGTTGGACTACTCTCGGAGGAGTGAGTGGTGGCCCGGCC 125
|||||

QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 126 GGCTTCGGCGCCACGAGGCGCGTCCAGCAGCAGTCTCAAGGAGAGTTCGAGGAG 185
|||||

QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
|||||
Db 186 GTCGACACGGTCTCAGCGCGCGCGCCACACCCACCCACCATGGTTCACCGGC 245
|||||

QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||
Db 246 GGCCACGGCTTCGTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 305
|||||

QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 306 GAGGTCCACGAGCGGAGGAGTCTCTCCGCCAGGGCTAAC 347
|||||

RESULT 10
BG840383
LOCUS
DEFINITION
sequence.
MEST12-H11.T7-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 550)

LOCUS AY104409 553 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PC0124784 mRNA sequence.
 ACCESSION AY104409
 VERSION AY104409.1 GI:21207487
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 553)
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 553)
 AUTHORS Coe,E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 FEATURES
 source
 Location/Qualifiers
 1..553
 /organism="Zea mays"
 /db_xref="taxon:637271"
 /clone_lib="PC0124784"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

BASE COUNT 128 a 173 c 162 g 90 t

Alignment Scores:
 Pred. No.: 1.11e-45 Length: 553
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 11 Gaps: 1

US-10-090-035-2 (1-93) x AY104409 (1-553)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 76 ATGGCTTACTACAGGAGGTGGACTACTCTCGAGGAGGTGAGGTGGCGCGGCC 135
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 136 GGCTTCGGCGGCCACGAGGCGGCTCCAGCAGCAGCTCTCAAGGAGATTCGAGGAG 195
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
 |||||
 Db 196 GTCGACAGGTCTTCACGCGCGCGGCCACCCACCACCATGGTCACACCGGC 255
 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspTyrCysThrGly 79
 |||||
 Db 256 GGCACAGCTCTGTTGGTGGCGAGACCGAGGTCTCGAGGAGGACATCAACCTGCACCGGC 315
 QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 316 GAGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 357

RESULT 12

LOCUS BQ619167 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ4E05_SK.abl Salt stressed Zea mays roots cDNA library Zea
 mays cDNA clone RN05EQ4E05_SK.abl similar to No homology, mRNA

sequence.
 ACCESSION BQ619167
 VERSION BQ619167.1 GI:21621161
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Wang,H. and Bohnert,H.J.
 TITLE Genomics of plant stress tolerance
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mark Fredricksen
 Department of Plant Biology
 University of Illinois
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 2172653473
 Email: bohnertlab@life.uiuc.edu.
 FEATURES
 source
 Location/Qualifiers
 1..648
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone="RN05EQ4E05_SK.abl"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /tissue_type="Roots"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150
 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

Alignment Scores:
 Pred. No.: 1.33e-45 Length: 648
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619167 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 82 ATGGCTTACTACAGGAGGTGGACTACTCTCGAGGAGGTGAGGTGGCGCGGCC 141
 QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 142 GGCTTCGGCGGCCACGAGGCGGCTCCAGCAGCAGCTCTCAAGGAGATTCGAGGAG 201
 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
 |||||
 Db 202 GTCGACAGGTCTTCACGCGCGCGGCCACCCACCACCATGGTCACACCGGC 261
 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspTyrCysThrGly 79
 |||||
 Db 262 GGCACAGCTCTGTTGGTGGCGAGACCGAGGTCTCGAGGAGGACATCAACCTGCACCGGC 321
 QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 322 GAGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 363

RESULT 13

LOCUS BQ619315 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ6C06_SK.abl Salt stressed Zea mays roots cDNA library Zea
 mays cDNA clone RN05EQ6C06_SK.abl similar to No homology, mRNA
 sequence.
 ACCESSION BQ619315
 VERSION BQ619315.1 GI:21621309
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

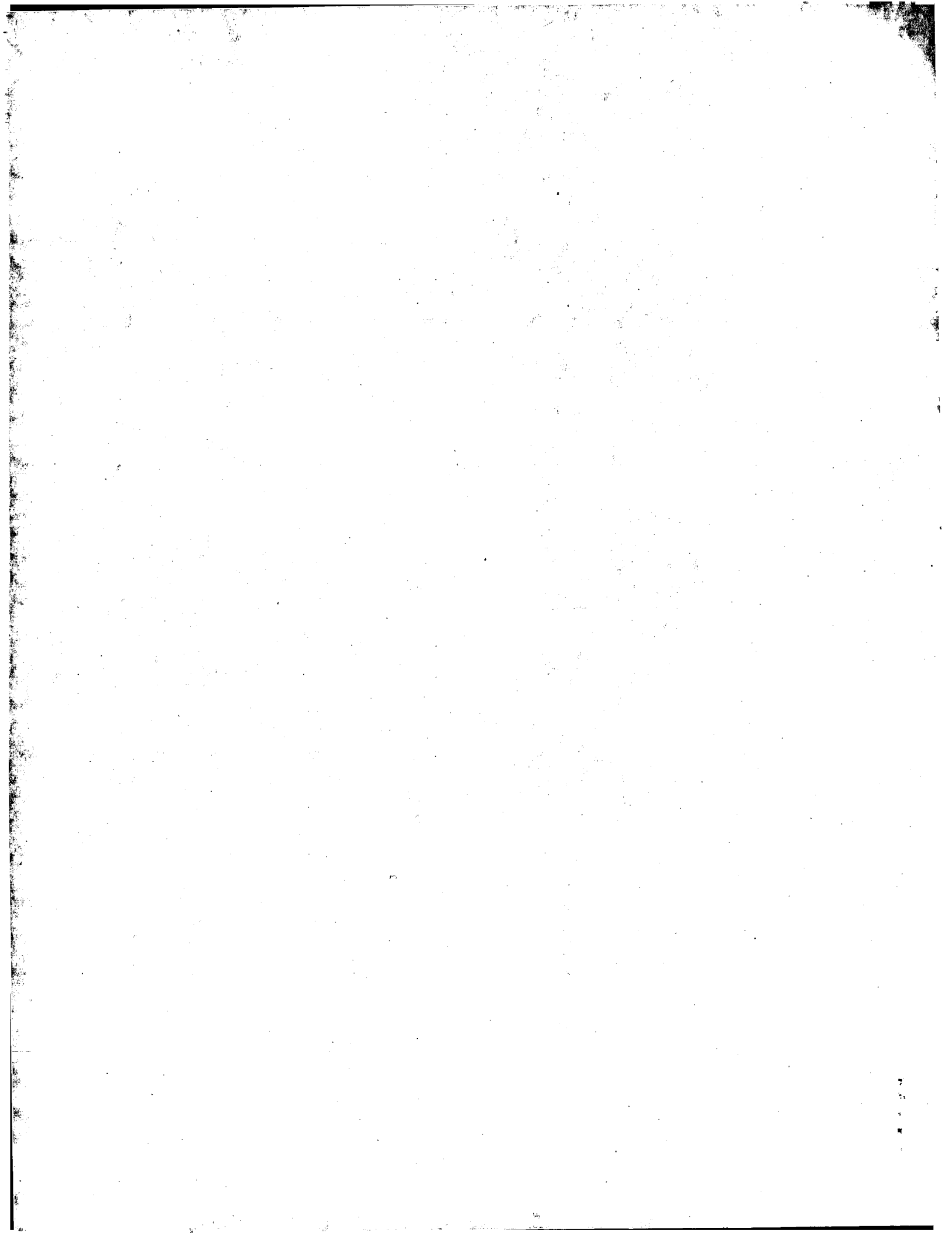
REFERENCE AUTHORS TITLE JOURNAL COMMENT	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu.	
	1 (bases 1 to 648) Wang, H. and Bohnert, H.J. Genomics of plant stress tolerance Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu.		Location/Qualifiers 1. 648 /organism="Zea mays" /db_xref="taxon:4577" /clone="RNOSEQ6D01.SK.ab1" /clone_lib="Salt stressed Zea mays roots cDNA library" /tissue_type="Roots" /dev_stage="2 weeks old" /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"	
	BASE COUNT ORIGIN		172 a 193 c 183 g 100 t	
FEATURES source	Alignment Scores:		Length: 648 Pred. No.: 93 Score: 499.50 Percent Similarity: 98.94% Best Local Similarity: 98.94% Query Match: 97.94% DB: 14	
	US-10-090-035-2 (1-93) x BQ619318 (1-648)			
	QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20 Db 82 ATGGCTTACTACCGAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCCGGCC 141 QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40 Db 142 GGCCTTCGGCCGACGAGCGCGCTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisGly 59 Db 202 GTCCACACGGTCTCACCGCGCGGCCCAACACCACCACCATGTCACACCGGC 261 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79 Db 262 GCCACCGCTTCGTGTGTCGCGAGACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321 QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93 Db 322 GAGGTCCACGAGCGGAGAGGCTTCCTCCGAGGGCTTAAC 363		648 bp mRNA linear EST 27-JUN-2002 RNOSEQ6D01.SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSEQ6D01.SK.ab1 similar to No homology, mRNA sequence. BQ619318 BQ619318.1 GI:21621312 EST. Zeas mays. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 648) Wang, H. and Bohnert, H.J. Genomics of plant stress tolerance Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473	
REFERENCE AUTHORS TITLE JOURNAL COMMENT	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu.	
	1 (bases 1 to 648) Wang, H. and Bohnert, H.J. Genomics of plant stress tolerance Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu.		Location/Qualifiers 1. 648 /organism="Zea mays" /db_xref="taxon:4577" /clone="RNOSEQ6D01.SK.ab1" /clone_lib="Salt stressed Zea mays roots cDNA library" /tissue_type="Roots" /dev_stage="2 weeks old" /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"	
	BASE COUNT ORIGIN		172 a 193 c 183 g 100 t	
FEATURES source	Alignment Scores:		Length: 648 Pred. No.: 93 Score: 499.50 Percent Similarity: 98.94% Best Local Similarity: 98.94% Query Match: 97.94% DB: 14	
	US-10-090-035-2 (1-93) x BQ619315 (1-648)			
	QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20 Db 82 ATGGCTTACTACCGAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCCGGCC 141 QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40 Db 142 GGCCTTCGGCCGACGAGCGCGCTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201 QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisGly 59 Db 202 GTCCACACGGTCTCACCGCGCGGCCCAACACCACCACCATGTCACACCGGC 261 QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79 Db 262 GCCACCGCTTCGTGTGTCGCGAGACGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321 QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93 Db 322 GAGGTCCACGAGCGGAGAGGCTTCCTCCGAGGGCTTAAC 363		648 bp mRNA linear EST 27-JUN-2002 RNOSEQ6D01.SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSEQ6D01.SK.ab1 similar to No homology, mRNA sequence. BQ619318 BQ619318.1 GI:21621312 EST. Zeas mays. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 648) Wang, H. and Bohnert, H.J. Genomics of plant stress tolerance Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473	

FEATURES Email: bohnertlab@life.uiuc.edu.
source Location/Qualifiers
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQE12_SK.ab1"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-2 (1-93) x BQ619337 (1-648)
QY 1 MetAlaTyrTyrGlnGluValAlaAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 82 ATGGCTTACTACCGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTCGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 142 GGCCTCGGCCGCCACGCGAGCGGGCTCCAGCAGCACCGTCGTCAAGGAGAAAGTTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisHisGly 59
Db 202 GTCGACACGGTCTCACGCGCGCGGCCAACCCACCACCACCACCATGTCACCACGGC 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCGTGGTGGCGAGACCAGGGTCGAGGAGGACATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGCAGGAGAGAGCTTCCTCGCCAGGGCTAAC 363

Search completed: June 16, 2003, 12:16:33
Job time : 1303.5 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 04:03:20 ; Search time 1684.8 Seconds
(without alignments)
9915.101 Million cell updates/sec

Title: US-10-090-035-3

Perfect score: 574

Sequence: 1 accacgcgtccgcacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	44	7.7	173699	8	AP003416
C 2	44	7.7	174218	2	AC091071
C 3	44	7.7	185095	8	AF003436
C 4	44	7.7	185481	2	AC091088
C 5	31	5.4	185133	2	AC119716
C 6	29	5.1	569	6	AX400832
C 7	29	5.1	1138	9	BC032756
C 8	29	5.1	1275	8	AY102542
C 9	29	5.1	1450	5	AF295407
C 10	29	5.1	2879	10	BC012959
C 11	29	5.1	3194	9	AK025519
C 12	29	5.1	30523	2	AC115592
C 13	29	5.1	112226	9	AC099486
C 14	29	5.1	172823	9	AC114323
C 15	28	4.9	410	8	AF467161
C 16	28	4.9	411	8	AF467129
C 17	28	4.9	411	8	AF467148
C 18	28	4.9	411	8	AF467156
C 19	28	4.9	881	3	AY044834
C 20	28	4.9	1436	9	BC015065
C 21	28	4.9	1635	9	BC010519
C 22	28	4.9	1849	9	BC008675
C 23	28	4.9	1853	3	AB062683
C 24	28	4.9	1865	9	BC012562
C 25	28	4.9	2115	10	BC007136
C 26	28	4.9	5928	6	AX346988
C 27	28	4.9	6593	6	AX345380
C 28	28	4.9	13919	6	AX356480
C 29	28	4.9	42499	8	AC004625
C 30	28	4.9	77722	9	AC093762
C 31	28	4.9	100652	2	AC116098
C 32	28	4.9	130010	2	AC025292
C 33	28	4.9	132000	2	AC116976
C 34	28	4.9	133888	2	AC010769
C 35	28	4.9	152472	2	AP000486
C 36	28	4.9	155354	2	AC107591
C 37	28	4.9	159658	9	AC104062
C 38	28	4.9	168145	9	AL159171
C 39	28	4.9	170022	2	AC021559
C 40	28	4.9	176827	9	AC022306
C 41	28	4.9	179191	2	AL608114
C 42	28	4.9	209764	10	AL645468
C 43	28	4.9	216250	9	AC099731
C 44	27	4.7	330	11	G41782
C 45	27	4.7	369	3	AF400199

ALIGNMENTS

RESULT 1
AP003416/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0456E05.
ACCESSION
AP003416
VERSION
AP003416.4
KEYWORDS
GI:20804922
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,
clone: P0456E05.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1

173699 bp DNA linear PLN 17-MAY-2002

CDS

```

complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/gene="P0456E05.7"
/notes="contains ESTs AU101665(R2414),D24709(R2594)
similar to Arabidopsis thaliana chromosome 5, At5g14420
unknown protein"
/codon_start=1
/protein_id="BAB92608.1"
/db_xref="GI:20804929"
/translation="MGQGSRSRDRSGHGRYGHSPFQQQWGGGGGGGGGGGYPY
GDADGGGYGAPQGGYAAAPAYQPPMPSPAPATOPPARAGGAKPRLDRYS
RIADYHSVDYTDALAAQAGLESSNLIVGIDTKSNEWLKFESHGRLHHISNAPN
YQAISIIIGOTLSKDEENLPCFGDATHDQVFCYFDLPCPCGSEALSRYRE
LPHLRAGTSPFADLEIEMANTIVEQSGQTHVLIITADQGVTSVDTASQLSQEQ
KVDAIVRASELPLSIVLVGVDGPDWMDKEFDNIPRAFDFQVNFSEIMKNMP
QSRKAARALMEIPQYKATVELGILGRSFKSPERVPLPPGGSHDAYSYGSKS
FKSPSTYSSSSPHYETAQSSPAVPSSTVDNQVCPICLVNPKDMAFGGCHQC
NPKAFERLADPKVYWCIDKDRKASNPQLQDCGAGDVSSIETH"
/gene="P0456E05.8"
51367..52887
/gene="P0456E05.8"
51367..52887
/notes="hypothetical protein
similar to salt-inducible protein"
/codon_start=1
/protein_id="BAB92609.1"
/db_xref="GI:20804930"
/translation="MAASRLRLKPLSLSKHQLRSLPEIDVEQAESPASSSIPID
PSLPLPLAVSHLSPPLPALPSVHASTPAALLRLRLRRHPRALADLHLLAAA
SDSAFRPDRLTSLASRAESRLPLSLRLLEVLFRPCPTDSDIFACPELPTF
RKAILAAAGDIPDAADALASRLAASPLPAEFYNILHALARLHREDTVEYFE
MTSHRHVPDYTTNLTINSRAEGVDAAWRFOEMRRWSCSPGTGVSNTLMRGFR
EGCKEGIKVAREMLQGLFGLSVASMEIMINGLRCGEPEPKAAEVFEFLVDGVPEG
FDCLDLVESLCVRNVEKAVVWELILERNWVSLGVPGVTFLECLMEKGLDKACQ
MGMWRAAEIVDPISCYIFEALCEAGRVTDNRNLQAKRKGFOADGFTYSIYVQG
FGROIRKEGAEVLDEMDAGYVNTIATNRLDLGLHMRGSMQLOKCSHRNAA"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"

```

gene

```

Query Match 7.7% Score 44; DB 8; Length 173699;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CDS

```

QY 103 ACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTG 146
|||||
Db 157564 ACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTG 157521

RESULT 2
AC091071 174218 bp DNA linear HTG 27-MAR-2001
LOCUS Oriza sativa (japonica cultivar-group) chromosome 1 clone P0456E05,
DEFINITION *** SEQUENCING IN PROGRESS ***, 3 ordered pieces.
AC091071
AC091071.1 GI:13450004
VERSION HTG: HTGS_PHASE2.
KEYWORDS Oriza sativa (japonica cultivar-group).
SOURCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Oryzae; Oryza.
1 (bases 1 to 174218)
Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
and Eun, M.-Y.
Unpublished
TITLE Oriza sativa PAC P0456E05 genomic sequence
JOURNAL 2 (bases 1 to 174218)
REFERENCE Hahn, J.-H., Eun, M.-Y. and Kim, H.-I.
AUTHORS Direct Submission
JOURNAL Submitted (27-MAR-2001) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology (NIAT), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@rda.go.kr,
Tel: 82-31-290-0309, Fax: 82-31-290-0308)
The PAC clone was provided by Japanese Rice Genome Research Program
of NIAR.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 53113: contig of 53113 bp in length
* 53114 53213: gap of unknown length
* 53214 172404: contig of 119191 bp in length
* 172405 172504: gap of unknown length
* 172505 174218: contig of 1714 bp in length.
FEATURES
source
1..174218
/organism="Oryza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0456E05"
BASE COUNT 49037 a 38377 c 37185 g 49417 t 202 others
ORIGIN
Query Match 7.7% Score 44; DB 2; Length 174218;
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 ACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTG 146
|||||
Db 16068 ACTACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTG 16111

RESULT 3
AP003436/c 185095 bp DNA linear PLN 17-APR-2002
LOCUS Oriza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION PAC clone:P0470A12.
AC003436
AC003436.2 GI:16197551
VERSION AP003436.2
KEYWORDS Oriza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
SOURCE clone:P0470A12.
ORGANISM Oriza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Oryzae; Oryza.
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oriza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12
Published Only in Database (2001)
2 (bases 1 to 185095)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (21-MAR-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Oct 17, 2001 this sequence version replaced gi:13430001.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGF. Protein homologues of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent

```


AARISVILKSVDSVGVDDVATRLKAVIKDNTLSLELYSVGILLSIKRQGHNVLPD
ADNTFRAIKALSODSGRWYDNTSAESSTFAAGIALEALSIVASLSDSVDSMIAYV
KNDIVLFDTKSYDDGTFFDEKHVDAAYKGPITTSASVVRGVTSPRAASGKINI
PGKILGLAKFFLIGLPGSACDFNQIEVFLVLSLPSKSVLTSKDLKVEVTVV
FGSAAPPLRVNLVQVLGSDSKVITTEKLOPDLDNNVHYLDIAPLKIDVGYLSVFE
ISLQEOEHETIYATGCTNTAEAFVTGLIKVDKAEIGISDNDAGTGVESQKIDLQKDT
VLSLANHOKRLSFLSPLGKTFKPHQVFLKLKHDSKVHFLVFPKSAQRFIVL
DTGLVKEVYILSGRVDLELAVDAAMENSFLRALGHIELDLPEAKPKAPPAQAVD
PESEKPKKEISHIFRSRPPKLSFAFTGLTLLPIVGLFGLMRLGVNLKNFPPL
PAPAFASLHFAGIGAVLLYVFLWIKLDTFTTKYLSPLGVLVFGVGHRLSYLSST
SAKQKTA"

gene
complement(Join(47978)..48061,48179..48239,48314..48378,
48478..48759,48837..48911,49020..49094,49242..49298,
49552..49625,49763..49864,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52039,52148..52561))
/clone="P0470A12.8"

CDS
complement(Join(47978)..48061,48179..48239,48314..48378,
48478..48759,48837..48911,49020..49094,49242..49298,
49552..49625,49763..49864,50111..50213,50317..50378,
50507..50561,50738..50841,50926..51048,51175..51219,
51967..52039,52148..52561))

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 146
|||||
Db 27507 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 27464
|||||

AC091088
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 1 clone P0470A12,
*** SEQUENCING IN PROGRESS ***; 4 ordered pieces.
AC091088
VERSION
AC091088.1 GI:13487944
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa (japonica cultivar-group).
ORGANISM
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 185481)
Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
and Eun, M.-Y.
Unpublished
2 (bases 1 to 185481)
Hahn, J.-H., Eun, M.-Y. and Kim, H.-I.
Direct Submission
Submitted (29-MAR-2001) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology(NAIST), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail:jhahn@rda.go.kr,
Tel:82-31-290-0309, Fax:82-31-290-0308)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11215: contig of 11215 bp in length
* 11316 11315: gap of unknown length
* 93352 93352: contig of 82037 bp in length
* 93453 93452: gap of unknown length
* 138783 138783: contig of 45331 bp in length
* 138784 138883: gap of unknown length
* 138884 185481: contig of 46598 bp in length.
Location/Qualifiers

FEATURES

source

1..185481
/organism="Oryza sativa (japonica cultivar-group)"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0470A12"
BASE COUNT 51439 a 41522 c 41493 g 50653 t 374 others
ORIGIN

Query Match

Best Local Similarity 7.7%; Score 44; DB 2; Length 185481;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 146
|||||

Db 156728 ACTACGAGGAGTGGACTACTGCTCGGAGGAGTGGAGTGGTG 156771
|||||

RESULT 5

AC119716

LOCUS

DEFINITION
Rattus norvegicus clone CH230-274116, *** SEQUENCING IN PROGRESS
***, 77 unordered pieces.
AC119716

ACCESSION

VERSION
AC119716.4 GI:21902555

KEYWORDS

HTG; HTGS_PHASE1.

SOURCE

Rattus norvegicus.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 185133)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Anaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbarie, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Chen, Z., Chowdhry, I., Christopoulos, C.,
Chen, G., Chen, R., Chen, Z., Coyle, M.D., Dathorne, S.R., David, R.,
Cleveland, C.D., Cox, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Davis, C., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Correll, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lodato, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission

TITLE


```

*      89844      91514: contig of 1671 bp in length

Query Match
Best Local Similarity 5.4%; Score 31; DB 2; Length 185133;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGCACAGC 31
|||||
Db 51879 ACCACGCGTCCGCCACGCGTCCGCACAGC 51909
|||||

RESULT 6
AX400832/c
LOCUS AX400832 569 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 508 from Patent WO0210453.
ACCESSION AX400832
VERSION AX400832.1 GI:21337012
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Eliashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 508 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
Source Location/Qualifiers
1..569
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="EMBL/GenBank Accession No. AI007824"
/translation="132 a 98 c 117 g 222 t"

BASE COUNT 132 a 98 c 117 g 222 t
ORIGIN
Query Match
Best Local Similarity 5.1%; Score 29; DB 6; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAATACTAAAAAATAAAAAAATAAAAAA 574
|||||
Db 73 GTAATACTAAAAAATAAAAAAATAAAAAA 45
|||||

RESULT 7
BC032756
LOCUS BC032756 1138 bp mRNA linear PRI 26-JUN-2002
DEFINITION Homo sapiens, H2A histone family, member C, clone MGC:44857
IMAGE:5581631, mRNA, complete cds.
ACCESSION BC032756
VERSION BC032756.1 GI:21595675
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL),
Sequencing Center (NISC),

*      89844      91514: contig of 1671 bp in length

Query Match
Best Local Similarity 5.4%; Score 31; DB 2; Length 185133;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGCACAGC 31
|||||
Db 51879 ACCACGCGTCCGCCACGCGTCCGCACAGC 51909
|||||

RESULT 6
AX400832/c
LOCUS AX400832 569 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 508 from Patent WO0210453.
ACCESSION AX400832
VERSION AX400832.1 GI:21337012
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Eliashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 508 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
Source Location/Qualifiers
1..569
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="EMBL/GenBank Accession No. AI007824"
/translation="132 a 98 c 117 g 222 t"

BASE COUNT 132 a 98 c 117 g 222 t
ORIGIN
Query Match
Best Local Similarity 5.1%; Score 29; DB 6; Length 569;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAATACTAAAAAATAAAAAAATAAAAAA 574
|||||
Db 73 GTAATACTAAAAAATAAAAAAATAAAAAA 45
|||||

RESULT 7
BC032756
LOCUS BC032756 1138 bp mRNA linear PRI 26-JUN-2002
DEFINITION Homo sapiens, H2A histone family, member C, clone MGC:44857
IMAGE:5581631, mRNA, complete cds.
ACCESSION BC032756
VERSION BC032756.1 GI:21595675
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL),
Sequencing Center (NISC),

Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Blakesley, R.W., Bouffard, G.G., Breese, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgel, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 69 Row: p Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15718718.

FEATURES
Source Location/Qualifiers
1..1138
/organism="Homo sapiens"
/db_xref="LOCUSID:8329"
/db_xref="taxon:9606"
/clone="MGC:44857 IMAGE:5581631"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_67"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
52..444
/codon_start=1
/product="H2A histone family, member C"
/protein_id="AAH32756.1"
/db_xref="GI:21595676"
/translation="MSGKQGGKQKARAKTSSBAGLQFPVGRVHLLRGNVAERV
GAGAPVLAIVLELTAELEAGNAADNKNKTRIPRLQLAKNDEELNKLGLKVT
IAQGVLPNIQAVILPKKTESHHKANG"

BASE COUNT 312 a 315 c 303 g 208 t
ORIGIN
Query Match
Best Local Similarity 5.1%; Score 29; DB 9; Length 1138;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCACGCGTCCGCCACGCGTCCGCACAG 30
|||||
Db 12 CCCACGCGTCCGCCACGCGTCCGCACAG 40
|||||

RESULT 8
AY102542
LOCUS AY102542 1275 bp mRNA linear PLN 13-AUG-2002
DEFINITION Arabidopsis thaliana hypothetical protein (At2g23370/F2686.2) mRNA,
complete cds.
ACCESSION AY102542
VERSION AY102542.1 GI:21805664
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1275)
Xiao, Y., Ishmael, N., Kumar, N., Redman, J., Riedmuller, S.,
Utterback, T., Whitelaw, C.A., Fraser, C.M. and Town, C.D.
Cloning and sequencing of full-length cDNAs for hypothetical genes
from chromosome 2 of Arabidopsis thaliana
Unpublished
2 (bases 1 to 1275)
Xiao, Y., Ishmael, N., Kumar, N., Redman, J., Riedmuller, S.,
Utterback, T., Whitelaw, C.A., Fraser, C.M. and Town, C.D.
Direct Submission
Submitted (08-MAY-2002) Plant Genomics, The Institute for Genomic
Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
Location/Qualifiers

```

```

1. .1275
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="2"
/notes="ecotype: Columbia"
1. .1275
/gene="At2g23370/F26B6.2"
29. .1051
/gene="At2g23370/F26B6.2"
/codon_start=1
/evidence="not_experimental"
/product="hypothetical protein"
/protein_id="AA076747.1"
/db_xref="GI:21805665"
/translation="MHRGAAGVGRKRRLLHLSRNFARSCSSSLSEHCFKIDIA
KAQPKHLMOLLNITARGKSTVSPKAGGLLPIPLVKMSPGSSIALLRWPTAPSS
MEMPVVEVKHGKWFELANNVDOFIHRIIVEDVSKPECSOEIENAAEGAKKLYSKG
DFASSRLMDLALYLLKRVGLFPDSLERKVIHRIENGDSVSLVATEFTYTKRGNFPGFA
RPFANAKVLLKLRNLAKDAARALKSSWTLCGRVEETIAQIAEWGEEOIAQYKER
VTGEGQRDIRDKRPMQAASLDEAAFLNLASLEGTWDESLEVAQCYKEAGLNDIAK
FVLYRD"
polyA_site
1183
/gene="At2g23370/F26B6.2"
/notes="alternative"
1240
/gene="At2g23370/F26B6.2"
/notes="alternative"
1251
/gene="At2g23370/F26B6.2"
1 others
BASE COUNT 386 a 232 c 308 g 348 t
ORIGIN
Query Match 5.1%; Score 29; DB 8; Length 1275;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 546 GTAATACTAAAAA...AAAAA 574
|||||
DB 1242 GTAATACTAAAAA...AAAAA 1270
|||||

RESULT 9
AF295407 1450 bp mRNA linear VRT 01-SEP-2001
LOCUS
DEFINITION
Danio rerio alcohol dehydrogenase (ADH) mRNA, complete cds.
ACCESSION
AF295407
VERSION
AF295407.1 GI:15428577
KEYWORDS
unpublished
SOURCE
Danio rerio.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 1450)
Tanguay,R.L. and Vasillou,V.
Identification and characterization of the zebrafish alcohol
dehydrogenase
unpublished
JOURNAL
2 (bases 1 to 1450)
Tanguay,R.L. and Vasillou,V.
Direct Submission
Submitted (15-AUG-2000) School of Pharmacy, University of Colorado
Health Sciences Center, 4200 East Ninth Ave, Box C238, Denver, CO
80262, USA
FEATURES
source
1. .1450
/organism="Danio rerio"
/db_xref="taxon:7955"
1. .1450
/gene="ADH"
49. .1182
/gene="ADH"
/codon_start=1

gene
CDS

/product="alcohol dehydrogenase"
/protein_id="AAK97853.1"
/db_xref="GI:15428578"
/translation="MATACKVVKRAAVAKPAPLMEEIEVAPPEGEIRKVIAT
GLCHTDLYHLVDGDKRGFFVVLGHESAGVVEGCVTDYKPGDKVPLPLSQCGKCK
FCCKPTNLCESSWATKYHDIMAEPTSRCTRGQTILQPMGTSTSEVTVINONAVAK
IDENAPLDNRVFLGCGITTYGAAVNTAGTSGVCAVFGIAGVGLAAVYMGCKNAGAS
RIFADVINEKKFKAKVFGVGTNDFNPKAFNKPISFLEIENGGVGVDFSECTGNIEV
MRSALESCAKGWSVVGVTNVDFNSAKPIQLIYGTWKVGLSGFGFKCKDSVPKLV
DYMGGKIMLDEFITHKMNLQVNDALNLMKTGGIIRTMTVSK"
BASE COUNT 419 a 316 c 356 g 359 t
ORIGIN
Query Match 5.1%; Score 29; DB 5; Length 1450;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CCACGGTCGCCGCCACGCGTCCGCACACG 31
|||||
DB 1 CCACGGTCGCCGCCACGCGTCCGCACACG 29
|||||

RESULT 10
BC012959 2879 bp mRNA linear ROD 07-AUG-2002
LOCUS
DEFINITION
Mus musculus, staufer (RNA-binding protein) homolog 1 (Drosophila),
clone MGC:13708 IMAGE:4188394, mRNA, complete cds.
ACCESSION
BC012959
VERSION
BC012959.1 GI:15277959
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2879)
Strausberg,R.
Direct Submission
Submitted (20-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 18 Row: j Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755673.
FEATURES
Location/Qualifiers
1. .2879
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:13708 IMAGE:4188394"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
290. .1759
/codon_start=1
/product="staufer (RNA-binding protein) homolog 1
(Drosophila)"
/protein_id="AAH12959.1"
CDS

```

```

/db_xref="GI:15277960"
/db_xref="LocusID:20853"
/translation="MYKVPDHSRQSTYSYCMRGAYPPRYFYFFPVPPLLYQVELS
VGGQPNKGKMRPVPKHDAPARALTQSEPLPERLEVNGREAEENLSEISQVF
EIALKNLPVNEVARESGPPHMKNEFTVSVYGEFVGEKGSKISKNKAARVLEQ
LRLPLPAVERVKPRIRKKSOPTKLTAPDYGGMNPISRLAIOQAKKEPEYM
LITERGLPRRFEVQVGVHHTAGVGNKVKARNAENMLETLGRKVPQAOAPK
ALKSEKTPVKPGDGKVTTFEPSPGDCGTSNKDEFRMPYLSHQQLPAGILWPV
EVAQAVGVSGHHTKDFPAAPNPAKATVAMIARELLYGTSPATETILKSNISGH
VPHGPRTPSEGLYLSRAQGVETKDFPPKNNKCVSLNCSQQPLVSHGIGKDV
ESCHDMAALNLIKLSELDQOSTMPRTGNGPVSACGR"
BASE COUNT      798 a 735 c 699 g 647 t
ORIGIN

Query Match      5.1%; Score 29; DB 10; Length 2879;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAACTACTAAAAA 574
|||||
Db 2821 GTAACTACTAAAAA 2849
|||||

RESULT 11
AK025519 3194 bp mRNA linear PRI 29-SEP-2000
LOCUS
DEFINITION Homo sapiens cDNA: FLJ21866 fis, clone HEP02379, highly similar to
ACCESSION AF061938 Homo sapiens staufen protein (STAU) mRNA.
VERSION AK025519
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens Hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
clone:HEP02379.
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Kawabata,A., Hiki,I., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,I., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3194)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1. .3194
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP02379"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"
misc_feature
1. .3194
/notes="highly similar to AF061938 Homo sapiens staufen
protein (STAU) mRNA"
BASE COUNT      894 a 787 c 732 g 781 t
ORIGIN

Query Match      5.1%; Score 29; DB 9; Length 3194;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAACTACTAAAAA 574
|||||
Db 3165 GTAACTACTAAAAA 3193
|||||

RESULT 12
AC115592/c 30523 bp DNA linear HTG 21-MAR-2002
LOCUS
DEFINITION Dictyostelium discoideum chromosome 2 map 109820-140341 strain AX4,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AC115592
VERSION HTG; HTGS PHASE2.
KEYWORDS Dictyostelium discoideum.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE 1 (bases 1 to 30523)
AUTHORS Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
Noegel,A.A.
Sequence and Analysis of Chromosome 2 of Dictyostelium
Unpublished
The Dictyostelium Genome Sequencing Consortium
2 (bases 1 to 30523)
Baumgart,C.
Direct Submission
Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
CDS predictions from GeneID may contain errors. Further Information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.unl-koeln.de/dictyostelium/project.shtml
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .30523
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/chromosomes="2"
/map="109820-140341"
697. .971
/notes="ORF_ID:dd_01553"
/pseudo
/codon_start=1
1886. .2778
/notes="ORF_ID:dd_01554"
/codon_start=1
/product="40S ribosomal protein S15A"
/protein_id="AAL92278.1"
/db_xref="GI:19569976"
/translation="MVRISLVNCLNSIVNAEROCKROVLVPPSSKVIIVKLEVMKK
RYIGFEIVDDHRSCKIVIDLGRINKGVISPRDVLDEIEKAWASVLLPSRQFGHI
VLITSLGINDHNEAKTRHTGKLLGFFY
complement(join(2504..2668,2759..3450,3569..3911))
/notes="ORF_ID:dd_01555"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAL92279.1"
/db_xref="GI:19569977"
/translation="MINELISKFKKVIYGFNCLVGEIFKSNKHLIGTVDGKII
VYLNDDEIEVLEVTGSGSQQLLELDITTFGSDVNSGNLVFSNHQILYRDTLN
GSITSITRLSIGDSLGLVTSIKPHQAINWKIPSVNQTLNSLIDSLIDAINPNIK

```



```

RESULT 13
AC099486
LOCUS 112226 bp DNA linear PRI 01-JAN-2002
DEFINITION Homo sapiens chromosome 5 clone CTD-2011G10, complete sequence.
ACCESSION AC099486
VERSION AC099486.2 GI:18030138
SOURCE HTG.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 112226)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 112226)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (01-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE
3 (bases 1 to 112226)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jan 1, 2002 this sequence version replaced gi:16930902.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
FEATURES
source
1. 112226
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2011G10"
BASE COUNT 36508 a 19376 c 19390 g 36952 t
ORIGIN
1. 112226
GTATATACTAAAAAAGAAAAAAGAAAAA 574
|||||
GTATATACTAAAAAAGAAAAAAGAAAAA 90193
|||||

Query Match 5.1%; Score 29; DB 9; Length 112226;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTATATACTAAAAAAGAAAAAAGAAAAA 574
|||||
GTATATACTAAAAAAGAAAAAAGAAAAA 90193
|||||

RESULT 14
AC114323/c
LOCUS 172823 bp DNA linear PRI 23-APR-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-93N19, complete sequence.
ACCESSION AC114323
VERSION AC114323.2 GI:20270116
SOURCE HTG.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 172823)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 172823)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
3 (bases 1 to 172823)

```

```

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 23, 2002 this sequence version replaced gi:19224972.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.3.
FEATURES
source
1. 172823
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-93N19"
BASE COUNT 56281 a 29916 c 30420 g 56206 t
ORIGIN
1. 172823
GTATATACTAAAAAAGAAAAAAGAAAAA 574
|||||
GTATATACTAAAAAAGAAAAAAGAAAAA 166759
|||||

Query Match 5.1%; Score 29; DB 9; Length 172823;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTATATACTAAAAAAGAAAAAAGAAAAA 574
|||||
GTATATACTAAAAAAGAAAAAAGAAAAA 166759
|||||

RESULT 15
AF467161
LOCUS 410 bp DNA linear PLN 15-JUL-2002
DEFINITION Lasthenia gracilis isolate calli53 trna-Lys (trnK) gene, intron;
and maturase K (matK) gene, partial cds; Chloroplast genes for
chloroplast products.
ACCESSION AF467161
VERSION AF467161.1 GI:21759599
KEYWORDS
SOURCE
ORGANISM Lasthenia gracilis.
Chloroplast Lasthenia gracilis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
Heleniidae; Lasthenia.
1 (bases 1 to 410)
Chan,R., Baldwin,B.G. and Ornduff,R.
Cryptic goldfields: A molecular phylogenetic re-investigation of
Lasthenia californica sensu lato and close relatives (Compositae:
Heliantheae sensu lato)
Am. J. Bot. 89 (7), 1103-1112 (2002)
2 (bases 1 to 410)
Chan,R., Baldwin,B.G. and Ornduff,R.
Direct Submission
JOURNAL Submitted (10-JAN-2002) Jepson Herbarium and Department of
Integrative Biology, University of California, 1001 Valley Life
Sciences Building # 2465, Berkeley, CA 94720-2465, USA
FEATURES
source
1. 410
/organism="Lasthenia gracilis"
/organelle="plastid:chloroplast"
/isolate="calli53"
/specimen_voucher="Neese and Painter 2552 (JEPS)"
/db_xref="taxon:182806"
<1..>410
/gene="trnK"
/note="CRNA-Lys"
<1..>410
/gene="trnK"
<1..175
/gene="matK"
<1..175
/gene="matK"
/codon_start=2

```

```
/transl_table=11
/product="maturase K"
/protein_id="AA077460.1"
/db_xref="GI:21759600"
/translation="RAFLKRFQSQLEEFFTEEEQVFSLTFFRVSSISRRLSRRIWY
LDIVCINDLANVE"
BASE COUNT      139 a      48 c      81 g      142 t
ORIGIN
Query Match      4.9%; Score 28; DB 8; Length 410;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      547 TAATACTAAAAA
Db      221 TAATACTAAAAA
```

Search completed: June 16, 2003, 07:56:05
Job time : 1698.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 06:35:31 ; Search time 39.9504 Seconds
(without alignments)
4406.275 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgcacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/lna/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/lna/backfilesi.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	4.5	212	4	US-09-280-116-50
2	26	4.5	638	4	US-08-998-416-981
3	26	4.5	717	4	US-08-998-416-981
4	26	4.5	792	3	US-08-950-720A-1
5	26	4.5	1021	4	US-09-280-116-85
6	26	4.5	1210	4	US-09-560-639-12
7	26	4.5	1384	4	US-09-372-422A-17
8	26	4.5	1448	2	US-08-942-218A-1
9	26	4.5	1504	4	US-09-280-116-1
10	26	4.5	2205	3	US-08-888-077A-41
11	26	4.5	2864	4	US-09-409-180A-2
12	26	4.5	3416	2	US-08-724-394A-15
13	26	4.5	3747	2	US-08-044-618-5
14	26	4.5	4880	3	US-09-031-563-1
15	26	4.5	4880	3	US-09-392-277-1
16	26	4.5	8519	4	US-09-261-907-1
17	26	4.5	162450	4	US-09-345-882-1
18	25	4.4	140	1	US-08-628-417-5
19	25	4.4	216	1	US-08-686-878A-34
20	25	4.4	216	4	US-09-175-928-34
21	25	4.4	240	1	US-08-628-417-6
22	25	4.4	330	4	US-09-078-294-24
23	25	4.4	493	4	US-09-280-116-166
24	25	4.4	493	4	US-09-392-184-20
25	25	4.4	508	4	US-09-280-116-47
26	25	4.4	601	4	US-09-336-536-74
27	25	4.4	728	4	US-09-091-097-5

```

28 25 4.4 749 4 US-09-257-583-12 Sequence 12, Appl
29 25 4.4 810 4 US-09-072-993C-5 Sequence 5, Appl
30 25 4.4 944 4 US-09-227-357-122 Sequence 122, App
31 25 4.4 1132 3 US-08-894-731-3 Sequence 3, Appl
32 25 4.4 1148 4 US-09-395-674B-7 Sequence 7, Appl
33 25 4.4 1164 2 US-08-794-796-1 Sequence 1, Appl
34 25 4.4 1284 4 US-09-286-529-6 Sequence 6, Appl
35 25 4.4 1353 4 US-08-913-014A-6 Sequence 44, Appl
36 25 4.4 1376 4 US-09-443-184-44 Sequence 2, Appl
37 25 4.4 1454 4 US-09-372-422A-19 Sequence 19, Appl
38 25 4.4 1619 2 US-08-991-946A-2 Sequence 2, Appl
39 25 4.4 1930 4 US-08-987-367-1 Sequence 1, Appl
40 25 4.4 2181 4 US-09-254-465A-7 Sequence 7, Appl
41 25 4.4 2181 4 US-09-254-465A-11 Sequence 11, Appl
42 25 4.4 2191 4 US-09-399-913-25 Sequence 25, Appl
43 25 4.4 2191 4 US-09-298-731-25 Sequence 25, Appl
44 25 4.4 2197 1 US-08-233-005-3 Sequence 3, Appl
45 25 4.4 2197 1 US-08-428-943-3 Sequence 3, Appl

```

ALIGNMENTS

RESULT 1

US-09-280-116-50
; Sequence 50, Application US/09280116A
; Patent No. 6331427

GENERAL INFORMATION:

APPLICANT: Robison, Keith E.

TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

FILE REFERENCE: 5800-24, 035800/176965

CURRENT APPLICATION NUMBER: US/09/280,116A

CURRENT FILING DATE: 1999-03-26

NUMBER OF SEQ ID NOS: 268

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 50

LENGTH: 212

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: zinc carboxypeptidases

US-09-280-116-50

Query Match 4.5%; Score 26; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCCAGCGTCGCCCGCCAGCGTCCGC 26

Db 19 ACCCAGCGTCGCCCGCCAGCGTCCGC 44

RESULT 2

US-08-998-416-981/c

; Sequence 981, Application US/08998416

; Patent No. 6239264

GENERAL INFORMATION:

APPLICANT: Philippsen, Peter

APPLICANT: Pohlmann, Rainer

APPLICANT: Steiner, Sabine

APPLICANT: Mohr, Christine

APPLICANT: Wendland, Jurgen

APPLICANT: Knechtle, Philipp

APPLICANT: Reibischung, Corinne

TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPYII

TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 1152

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6239264artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: No. 6239264th Carolina

COUNTRY: USA

ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998,416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 981:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 638 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: PAG1604RP
 US-08-998-416-981

Query Match 4.5%; Score 26; DB 4; Length 638;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 ATACTAAAAAAAAAAAAAAAAAAAAA 574
 ||||||||||||||||||||
 Db 624 ATACTAAAAAAAAAAAAAAAAAAAAA 599

RESULT 3
 US-08-998-416-904/c
 ; Sequence 904, Application US/08998416
 ; Patent No. 6239264
 ; GENERAL INFORMATION:
 ; APPLICANT: Philippsen, Peter
 ; APPLICANT: Pohlmann, Rainer
 ; APPLICANT: Steiner, Sabine
 ; APPLICANT: Mohr, Christine
 ; APPLICANT: Wendland, Jurgen
 ; APPLICANT: Knechtle, Philipp
 ; APPLICANT: Rebeschung, Corinne
 ; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
 ; TITLE OF INVENTION: AND USES THEREOF
 ; NUMBER OF SEQUENCES: 1152
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6239264artis Corporation
 ; STREET: 3054 Cornwallis Road
 ; CITY: Research Triangle Park
 ; STATE: No. 6239264th Carolina
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/998,416
 ; FILING DATE: 24-DEC-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 904:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 717 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: PAG1563UP
 US-08-998-416-904

Query Match 4.5%; Score 26; DB 4; Length 717;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 549 ATACTAAAAAAAAAAAAAAAAAAAAA 574
 ||||||||||||||||||||
 Db 171 ATACTAAAAAAAAAAAAAAAAAAAAA 146

RESULT 4
 US-08-950-720A-1/c
 ; Sequence 1, Application US/08950720A
 ; Patent No. 6046028
 ; GENERAL INFORMATION:
 ; APPLICANT: Conklin, Darrell C.
 ; APPLICANT: Lofton-Day, Catherine E.
 ; APPLICANT: Lok, Si
 ; APPLICANT: Jaspers, Stephen R.
 ; TITLE OF INVENTION: INSULIN HOMOLOG
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 1201 Eastlake Avenue East
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/950,720A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sawislak, Deborah A
 ; REGISTRATION NUMBER: 37,438
 ; REFERENCE/DOCKET NUMBER: 96-09
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-442-6672
 ; TELEFAX: 206-442-6678
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 792 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA

;
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 77...481
; OTHER INFORMATION:
US-08-950-720A-1

Query Match 4.5%; Score 26; DB 3; Length 792;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
|||||
DB 42 ACCACGCGTCCGCCACGCGTCCGC 17

RESULT 5

US-09-280-116-85/c
; Sequence 85, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: aspartyl proteases
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1021)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-85

Query Match 4.5%; Score 26; DB 4; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
|||||
DB 38 ACCACGCGTCCGCCACGCGTCCGC 13

RESULT 6

US-09-560-639-12
; Sequence 12, Application US/09560639
; Patent No. 6323334
; GENERAL INFORMATION:
; APPLICANT: Kingsbury, G.
; APPLICANT: Leiby, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/560,639
; CURRENT FILING DATE: 2000-04-28
; EARLIER APPLICATION NUMBER: 60/155,862
; EARLIER FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(557)
US-09-560-639-12

Query Match 4.5%; Score 26; DB 4; Length 1210;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
|||||
DB 5 ACCACGCGTCCGCCACGCGTCCGC 30

RESULT 7

US-09-372-422A-17
; Sequence 17, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)...(959)
US-09-372-422A-17

Query Match 4.5%; Score 26; DB 4; Length 1384;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCCGCCACGCGTCCGC 26
|||||
DB 10 ACCACGCGTCCGCCACGCGTCCGC 35

RESULT 8

US-08-942-218A-1/c
; Sequence 1, Application US/08942218A
; Patent No. 5935835
; GENERAL INFORMATION:
; APPLICANT: McCard-Roshak, Amy
; APPLICANT: Marshall, Lisa
; TITLE OF INVENTION: Human Myt-1 Kinase Clone
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/942,218A
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,389
; FILING DATE: 11-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031

```

; REFERENCE/DOCKET NUMBER: ATG-50027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-942-218A-1

Query Match
Best Local Similarity 4.5%; Score 26; DB 2; Length 1448;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGCGTCCGC 26
Db 46 ACCACGGCTCCGCCACGCGTCCGC 21

RESULT 9
US-09-280-116-1/c
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
; US-09-280-116-1

Query Match
Best Local Similarity 4.5%; Score 26; DB 4; Length 1504;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGCGTCCGC 26
Db 64 ACCACGGCTCCGCCACGCGTCCGC 39

RESULT 10
US-08-888-077A-41
; Sequence 41, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2205
; OTHER INFORMATION: /note= "mutTMI-TM2"
; US-08-888-077A-41

Query Match
Best Local Similarity 4.5%; Score 26; DB 3; Length 2205;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGCGTCCGC 26
Db 15 ACCACGGCTCCGCCACGCGTCCGC 40

RESULT 11
US-09-409-180A-2
; Sequence 2, Application US/09409180A
; Patent No. 6444802
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: White, David
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 22196, A No. 644802el Human Aminopeptidase
; FILE REFERENCE: 5800-59
; CURRENT APPLICATION NUMBER: US/09/409,180A
; CURRENT FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (300)...(2414)
; US-09-409-180A-2

Query Match
Best Local Similarity 4.5%; Score 26; DB 4; Length 2864;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGGCTCCGCCACGCGTCCGC 26
Db 250 ACCACGGCTCCGCCACGCGTCCGC 275

RESULT 12
US-08-724-394A-15/c
; Sequence 15, Application US/08724394A
; Patent No. 5872237
```

;; GENERAL INFORMATION:
;; APPLICANT: Feder, John N.
;; APPLICANT: Krommal, Gregory S.
;; APPLICANT: Lauer, Peter M.
;; APPLICANT: Ruddy, David A.
;; APPLICANT: Thomas, Winston
;; APPLICANT: Tsuchihashi, Zenta
;; APPLICANT: Wolff, Roger K.
;; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
;; TITLE OF INVENTION: Sequences and Antibodies Thereto
;; NUMBER OF SEQUENCES: 31
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/724,394A
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fitts, Renee A.
;; REGISTRATION NUMBER: 35,136
;; REFERENCE/DOCKET NUMBER: 017957-000100
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-576-0200
;; TELEFAX: 415-576-0300
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3416 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1..3416
;; OTHER INFORMATION: /note= "cDNA 44"
US-08-724-394A-15

Query Match 4.5%; Score 26; DB 2; Length 3416;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAGCGTCCGCCACGGTCCGC 26
Db 63 ACCCAGCGTCCGCCACGGTCCGC 38

RESULT 13
US-08-044-618-5
; Sequence 5, Application US/08044618
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/044,618
;; FILING DATE: 19930406
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/257,696
;; FILING DATE: 14-OCT-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOX, SAMUEL L
;; REGISTRATION NUMBER: 30,353
;; REFERENCE/DOCKET NUMBER: 0654, 0490001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)466-0800
;; TELEFAX: (202)833-8716
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3747 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: both
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-044-618-5

Query Match 4.5%; Score 26; DB 1; Length 3747;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 549 ATACTAAAAAATAAAAAAAAAAAAAA 574
Db 3688 ATACTAAAAAATAAAAAAAAAAAAAA 3713

RESULT 14
US-09-031-563-1
; Sequence 1, Application US/09031563A
; Patent No. 6022708
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Rozenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: FI272
; CURRENT APPLICATION NUMBER: US/09/031,563A
; CURRENT FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 4880
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unknown N
; LOCATION: 4160, 4243, 4361
; OTHER INFORMATION: unknown
US-09-031-563-1

Query Match 4.5%; Score 26; DB 3; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAGCGTCCGCCACGGTCCGC 26
Db 33 ACCCAGCGTCCGCCACGGTCCGC 58

RESULT 15
US-09-392-277-1
; Sequence 1, Application US/09392277A
; Patent No. 6451977
; GENERAL INFORMATION:

```

; APPLICANT: Frederic de Sauvage
; APPLICANT: Airon Rosenthal
; FILE OF INVENTION: Fused
; FILE REFERENCE: P1272R1P1
; CURRENT APPLICATION NUMBER: US/09/392,277A
; CURRENT FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US 09/258,000
; EARLIER FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: US 60/076,072
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 28
; SEQ ID NO 1
; LENGTH: 4880
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unknown N
; LOCATION: 4160, 4243, 4361
; OTHER INFORMATION: unknown
US-09-392-277-1

Query Match      4.5%; Score 26; DB 4; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ACCGCGGTGCGCGCCGCGTCCGC 26
         |||||||
Db      33  ACCGCGGTGCGCGCCGCGTCCGC 58

Search completed: June 16, 2003, 08:47:04
Job time : 42.9504 secs

```

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 03:58:29 ; Search time 188.731 Seconds
(without alignments)
6849.143 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgcgcgcacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	31	5.4	908	22	AAF97898
2	29	5.1	569	24	ABK62601
3	29	5.1	1926	22	AAF81796
4	28	4.9	296	22	AAK57062
5	28	4.9	401	22	AAI87312
6	28	4.9	532	20	AAK22224
7	28	4.9	1592	24	ABQ54307
8	28	4.9	2874	22	AAI57944
9	28	4.9	4929	22	AAK51691
10	28	4.9	5928	24	ABL34086
11	28	4.9	6593	24	ABL32478
12	28	4.9	13919	24	ABL22305
13	27	4.7	51	22	AAI30125
14	27	4.7	233	22	AAH33064
15	27	4.7	380	22	AAI93735
16	27	4.7	391	22	AAI88702
17	27	4.7	393	22	AAI88879
18	27	4.7	400	21	AAK98255
19	27	4.7	451	22	AAK88891
20	27	4.7	451	22	AAI57612
21	27	4.7	471	24	ABL67877
22	27	4.7	474	22	ABA57396
23	27	4.7	474	22	ABA26916
24	27	4.7	474	22	AAK05432
25	27	4.7	474	22	AAK31031
26	27	4.7	474	22	ABQ25591
27	27	4.7	474	22	ABQ25591
28	27	4.7	474	22	ABQ25591
29	27	4.7	474	22	ABQ25591
30	27	4.7	474	22	ABQ25591
31	27	4.7	474	22	ABQ25591
32	27	4.7	474	22	ABQ25591
33	27	4.7	474	22	ABQ25591
34	27	4.7	474	22	ABQ25591
35	27	4.7	474	22	ABQ25591
36	27	4.7	474	22	ABQ25591
37	27	4.7	474	22	ABQ25591
38	27	4.7	474	22	ABQ25591
39	27	4.7	474	22	ABQ25591
40	27	4.7	474	22	ABQ25591
41	27	4.7	474	22	ABQ25591
42	27	4.7	474	22	ABQ25591
43	27	4.7	474	22	ABQ25591
44	27	4.7	474	22	ABQ25591
45	27	4.7	474	22	ABQ25591

10	28	4.9	4929	22	AAI58784	Human polynucleoti
c 11	28	4.9	5928	24	ABL34086	Human immune syste
c 12	28	4.9	6593	24	ABL32478	Human immune syste
c 13	28	4.9	13919	24	ABL22305	Chemically treated
c 14	27	4.7	51	22	AAI30125	Human SNP oligonuc
15	27	4.7	233	22	AAH33064	Human colon cancer
16	27	4.7	380	22	AAI93735	Human polynucleoti
17	27	4.7	391	22	AAI88702	Human polynucleoti
18	27	4.7	393	22	AAI88879	Human polynucleoti
19	27	4.7	400	21	AAK98255	Human colon cancer
20	27	4.7	451	22	AAK88891	Human polynucleoti
21	27	4.7	451	22	AAI57612	Human digestive sy
22	27	4.7	471	24	ABL67877	Human colorectal c
23	27	4.7	474	22	ABA57396	Ovary cancer relat
24	27	4.7	474	22	ABA26916	Human foetal liver
25	27	4.7	474	22	AAK05432	Probe #5382 for ge
26	27	4.7	474	22	AAK31031	Human brain expres
27	27	4.7	474	22	ABQ25591	Human bone marrow
28	27	4.7	474	22	ABQ25591	Probe #5627 used t
c 29	27	4.7	474	22	ABQ25591	Human genome-deriv
30	27	4.7	474	22	ABQ25591	cDNA encoding SRT
31	27	4.7	474	22	ABQ25591	Human polynucleoti
32	27	4.7	474	22	ABQ25591	Human polynucleoti
33	27	4.7	474	22	ABQ25591	Human immune/haema
34	27	4.7	474	22	ABQ25591	Human polynucleoti
c 35	27	4.7	474	22	ABQ25591	Human prostate exp
c 36	27	4.7	474	22	ABQ25591	Human colon cancer
c 37	27	4.7	474	22	ABQ25591	DNA encoding novel
38	27	4.7	474	22	ABQ25591	Oligonucleotide fo
39	27	4.7	474	22	ABQ25591	Oligonucleotide fo
40	27	4.7	474	22	ABQ25591	Human colon cancer
41	27	4.7	474	22	ABQ25591	Human polynucleoti
42	27	4.7	474	22	ABQ25591	Human TANGO 130 cD
43	27	4.7	474	22	ABQ25591	Human polynucleoti
44	27	4.7	474	22	ABQ25591	Human secreted pro
45	27	4.7	474	22	ABQ25591	Corn clone 7002148
						Human polynucleoti

ALIGNMENTS

RESULT 1
AAF97898
ID AAF97898 standard; cDNA; 908 BP.
AC AAF97898;
XX
DT 01-JUN-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 25.
XX
KW Human; secreted protein; immunomodulatory; antisclerotic;
KW dermatological; anti-inflamatory; anti-HIV; cytostatic; cardiant;
KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW neurotropic; anticonvulsant; antialzheimers; antiparkinsonian;
KW antimicrobial; vulnerrary; vaccine; gene therapy; cancer;
KW protein coordinate data; infection; ss.
XX
OS Homo sapiens.
XX
PN WO200121658-A1.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-US26013.
XX
PR 24-SEP-1999; 99US-0155709.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
PI Young PE, Wei P, Florence KA;

WPI; 2001-235311/24.

Nucleic acids encoding 32 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

Claim 1; Page 730; 890pp; English.

The present sequence encodes one of 32 novel human secreted polypeptides. The nucleic acid molecules and polypeptides they encode may be used in the prevention, diagnosis and treatment of diseases such as immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus and human immuno-deficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft neovascularisation and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease), infectious diseases and/or for promoting wound healing, regeneration and/or chemotaxis. The nucleic acid molecules may be used to produce the secreted polypeptides. They may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of their expression and activity.

Sequence 908 BP; 242 A; 290 C; 197 G; 175 T; 4 other;

Query Match 5.4%; Score 31; DB 22; Length 908;
Best Local Similarity 100.0%; Pred. NO. 0.0039;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACCCACGCGTCCGCCACGCGTCCGCACAGC 31
19 ACCCACGCGTCCGCCACGCGTCCGCACAGC 49

RESULT 2
ID ABRK62601/C
AC ABRK62601 standard; cDNA; 569 BP.
AC ABRK62601;
DT DT
XX 18-JUN-2002 (first entry)
DE
DE
KW Rat sequence differentially expressed in response to a hepatotoxin #508.
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differentially expression; centrilobular necrosis; steatosis.
OS Rattus norvegicus.
PN WO200210453-A2.
PD 07-FEB-2002.
XX 30-JUL-2001; 2001WO-US23872.
XX 31-JUL-2000; 2000US-222040P.
XX 02-NOV-2000; 2000US-244880P.
XX 11-MAY-2001; 2001US-290029P.
XX 15-MAY-2001; 2001US-290645P.
XX 22-MAY-2001; 2001US-292336P.
XX 06-JUN-2001; 2001US-295798P.
XX 13-JUN-2001; 2001US-297457P.
XX 19-JUN-2001; 2001US-298884P.
XX 09-JUL-2001; 2001US-303459P.
(GENE-) GENE LOGIC INC.
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

```

XX PD 22-FEB-2001.
XX PF
XX PR 16-AUG-2000; 2000WO-US22325.
XX PR 17-AUG-1999; 99US-0149182.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PR Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
XX PR Birse CE, Young PE, Komatsculis GA, Moore PA, Soppet DR;
XX PR WPI; 2001-147550/15.
XX PR P-PSDB; AAB74742.
XX PR
XX PT Nucleic acids encoding 25 human secreted polypeptides, useful for
XX PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX PT disease and diabetic retinopathy -
XX PR
XX PS Claim 1; Page 446; 485pp; English.
XX CC
XX CC AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
XX CC to AAB74772. Human secreted proteins can have activities based on the
XX CC tissues and cells they are expressed in. Example of activities include:
XX CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
XX CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiac;
XX CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
XX CC anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial; and
XX CC vulnerary. Human secreted proteins can be used in gene therapy and
XX CC vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins
XX CC (PEP1) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. For example, NAM1
XX CC and PEP1 may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patients genome
XX CC that affect the activity of proteins or expressing inactive proteins or
XX CC to supplement the patients own production of polypeptides. Disorders that
XX CC may be prevented, diagnosed and/or treated include immune disorders,
XX CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
XX CC angogenic disorders, neurological disorders, infectious diseases and/or
XX CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
XX CC AAF81786 and AAB74732 represent sequences used in the exemplification of
XX CC the present invention.
XX SQ Sequence 1926 BP; 519 A; 429 C; 430 G; 548 T; 0 other;

Query Match 5.18; Score 29; DB 22; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 GTAATACATAAAAAAAAAAAAAAAAAAAAAA 574
Db 1881 GTATATACATAAAAAAAAAAAAAAAAAAAAAA 1909

RESULT 4
AAK57062
ID AAK57062 standard; cDNA; 296 BP.
XX AC
XX AC AAK57062;
XX DT
XX DT 06-NOV-2001 (first entry)
XX DE
XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2122.
XX KW
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX OS
XX OS Homo sapiens.
XX OS WO200157182-A2.
XX PN
XX PN 09-AUG-2001.
XX PD
XX PD

```

29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 13-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 P-PSDB; AAM84281.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX Claim 1; SEQ ID NO 2122; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC the nucleic acids may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells: AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX Sequence 296 BP; 110 A; 66 C; 57 G; 61 T; 2 other;
 SQ

Query Match 4.9%; Score 28; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 0.074; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 0;
 QY 2 CCCACGCGTCCGCCACGCGTCCGCACA 29
 DB 17 CCCACGCGTCCGCCACGCGTCCGCACA 44
 RESULT 5
 AAI87312
 ID AAI87312 standard; cDNA; 401 BP.
 XX AAI87312;
 AC AAI87312;
 DT 06-NOV-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 7372.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX Homo sapiens.
 OS WO200164835-A2.
 PN 07-SEP-2001.
 XX 26-FEB-2001; 2001WO-US04927.
 XX 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR P-PSDB; AAO7381.
 XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX Claim 1; SEQ ID NO 7372; 1399pp + Sequence Listing; English.
 PS

XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.

XX Sequence 401 BP; 92 A; 110 C; 78 G; 121 T; 0 other;

Query Match 4.9%; Score 28; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCACGGCTCCGCCACGGCTCCGCACA 29
 |||||
 DB 1 CCCACGGCTCCGCCACGGCTCCGCACA 28

RESULT 6

AAAX22224
 ID AAAX22224 standard; DNA; 532 BP.

AC AAAX22224;

DT 18-MAY-1999 (first entry)

DE Human secreted protein gene 14 clone HHEPT60.

XX Human; secreted protein; gene therapy; protein therapy; cancer; weight;
 KW tumour; chromosome mapping; forensic; haematological disease; allergy;
 KW inflammation; cell proliferation; viral infection; wound healing;
 KW modulation; appetite; behaviour; food additive; preservative; ss.

OS Homo sapiens.

PN WO9903990-A1.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US14613.

PR 18-AUG-1997; 97US-0056361.
 PR 16-JUL-1997; 97US-0052661.
 PR 16-JUL-1997; 97US-0052870.
 PR 16-JUL-1997; 97US-0052871.
 PR 16-JUL-1997; 97US-0052872.
 PR 16-JUL-1997; 97US-0052873.
 PR 16-JUL-1997; 97US-0052874.
 PR 16-JUL-1997; 97US-0052875.
 PR 22-JUL-1997; 97US-0053440.
 PR 22-JUL-1997; 97US-0053441.
 PR 22-JUL-1997; 97US-0053442.
 PR 18-AUG-1997; 97US-0055683.
 PR 18-AUG-1997; 97US-0055724.
 PR 18-AUG-1997; 97US-0055725.
 PR 18-AUG-1997; 97US-0055726.
 PR 18-AUG-1997; 97US-0055946.
 PR 18-AUG-1997; 97US-0055952.
 PR 18-AUG-1997; 97US-0055985.
 PR 18-AUG-1997; 97US-0055989.
 PR 18-AUG-1997; 97US-0056359.

(HUMA-) HUMAN GENOME SCI INC.

XX

PI

PI Duan R, Feng P, Ferrie AM, Florence KA, Fouad J;
 PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE;
 XX Yu G;

DR WPI: 1999-132234/11.

XX P-PSDB; AAY01396.

PT New nucleic acids encoding secreted human proteins - potentially
 PT useful for treating and diagnosing diseases and identifying specific
 PT binding agents

PS Claim 4; Page 176; 251pp; English.

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX WO200200677-A1.
 XX 03-JAN-2002.
 XX 07-JUN-2001; 2001WO-US18569.
 XX 07-JUN-2000; 2000US-209467P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI: 2002-147878/19.
 XX P-PSDB; ABP41230.
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 XX useful in the prevention, treatment and diagnosis of cancer (e.g.
 XX ovarian cancer), immune disorders, cardiovascular disorders and
 XX neurological diseases -
 XX Claim 1; SEQ ID NO 187; 2922pp; English.
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 XX ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
 XX encompasses polypeptides 90% identical and polynucleotides 95% identical
 XX to the sequences of the invention. The invention additionally relates to
 XX recombinant vectors and host cells comprising human ovarian antigen
 XX polynucleotides, antibodies against human ovarian antigens, and the use
 XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
 XX treating, prognosing or preventing various ovarian and/or breast-related
 XX disorders. Such conditions include ovarian cancer and breast cancer, and
 XX metastatic tumours of ovarian or breast origin, reproductive system
 XX disorders (e.g., infertility, disorders of pregnancy, anovulation
 XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 XX vaginitis), immune disorders (e.g., congenital and acquired
 XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
 XX respiratory disorders, neurological disorders, gastrointestinal disorders
 XX and urinary system disorders. Ovarian antigen polypeptides and
 XX polynucleotides may also be used in screening for compounds which
 XX modulate ovarian antigen expression or activity. The polynucleotides may
 XX further be used for gene therapy, chromosome mapping, in the
 XX identification of individuals and in forensic analysis, and the
 XX polypeptides may be used as food additives or to prepare antibodies
 XX useful in disease diagnosis, drug targeting and phenotyping. The present
 XX sequence represents cDNA encoding a human ovarian antigen of the
 XX invention.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1592 BP; 458 A; 322 C; 338 G; 474 T; 0 other;
 XX
 XX Query Match 4.9%; Score 28; DB 24; Length 1592;
 XX Best Local Similarity 100.0%; Pred. No. 0.054;
 XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 ACCACGGTCCGCCACGGTCCGCAC 28
 XX |
 XX Db 4 ACCACGGTCCGCCACGGTCCGCAC 31
 XX
 XX RESULT 8
 XX AAI57944
 XX ID AAI57944 standard; cDNA; 2674 BP.
 XX AC AAI57944;
 XX XX
 XX DT 22-OCT-2001 (first entry)
 XX XX

DE Human polynucleotide SEQ ID NO 147.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM38788.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 XX Claim 1; SEQ ID NO 147; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 XX the encoded polypeptides (AAM38642-AAI62213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression,
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukaemia and
 XX C.N.S disorders.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification.
 XX Sequence 2674 BP; 906 A; 529 C; 580 G; 659 T; 0 other;
 XX
 XX Query Match 4.9%; Score 28; DB 22; Length 2674;
 XX Best Local Similarity 100.0%; Pred. No. 0.049;
 XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 ACCACGGTCCGCCACGGTCCGCAC 28
 XX |
 XX Db 21 ACCACGGTCCGCCACGGTCCGCAC 48
 XX
 XX RESULT 9
 XX AAK51691
 XX ID AAK51691 standard; cDNA; 4929 BP.
 XX XX
 XX AC AAK51691;

XX ABL34086;
 XX 26-MAR-2002 (first entry)
 XX Human immune system associated gene SEQ ID NO: 2059.
 XX Human: immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX Claim 1; SEQ ID NO 2059; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX Sequence 5928 BP; 1649 A; 110 C; 1223 G; 2946 T; 0 other;
 SQ
 Query Match 4.9%; Score 28; DB 24; Length 5928;
 Best Local Similarity 100.0%; Pred. No. 0.043;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 574
 DB 5865 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 5838
 RESULT 12
 ABL32478/c
 ID ABL32478 standard; DNA; 6593 BP.
 XX ABL32478;
 XX 26-MAR-2002 (first entry)
 XX Human immune system associated gene SEQ ID NO: 451.
 XX Human: immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;

KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX Claim 1; SEQ ID NO 451; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX Sequence 6593 BP; 1732 A; 121 C; 1448 G; 3292 T; 0 other;
 SQ
 Query Match 4.9%; Score 28; DB 24; Length 6593;
 Best Local Similarity 100.0%; Pred. No. 0.042;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 574
 DB 1002 TAATCTATAAAAAAAAAAAAAAAAAAAAAA 975
 RESULT 13
 ABL92305/c
 ID ABL92305 standard; DNA; 13919 BP.
 XX ABL92305;
 XX 01-JUL-2002 (first entry)
 XX Chemically treated DNA repair gene fragment complementary to#57.
 XX DNA repair; cytosine methylation; PMS2L1; PMS2L2; PMS2L3;
 KW PMS2; L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPP1; RFC4;
 KW DDT1L; FANCB; XRCC8; ataxia telangiectasia; aging; Bloom's syndrome;
 KW Cockayne syndrome; Nijmegen breakage syndrome; Werner syndrome;
 KW immunodeficiency; trichothiodystrophy; Fanconi's anaemia; solid tumour;
 KW cancer; ds.
 XX Unidentified.
 OS
 XX WO200181622-A2.
 XX 01-NOV-2001.
 XX 06-APR-2001; 2001WO-EP03972.

PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-034446/04.
 DR New nucleic acid derived from genes associated with DNA repair, useful
 XX for diagnosis, e.g. of ataxia telangiectasia, by determination of
 PT cytosine methylation.
 PT Claim 1; SEQ ID NO 114; 25pp + sequence listing; English.
 PS
 CC The invention relates to nucleic acids containing a sequence of at least
 CC 18 nucleotides of chemically treated DNA of genes associated with DNA
 CC repair, and their complements. The invention also relates to nucleic
 CC acids comprising at least 18 base pairs of the chemically pretreated DNA
 CC of genes associated with DNA repair selected from PMS2L1, PMS2L12,
 CC PMS2L2, PMS2L3, PMS2, L4, PMS2L5, PMS2L6, MGMT, MSH2, NUDT1, TDG, INPPL1,
 CC RFO4, DDIRIL, FANCB, or XRCC8. Nucleic acids of the invention and related
 CC oligomers, are useful for diagnosis of diseases associated with gene
 CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
 CC Cockayne syndrome, Mijmegen breakage syndrome or Werner syndrome,
 CC immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours
 CC and cancer, particularly by determining status of cytosine methylation
 CC and/or by detecting single-nucleotide polymorphisms. Determination of
 CC individual methylation patterns may allow development of individualised
 CC therapies. The sequences given in records ABL92192-ABL92335 represent
 CC chemically pre-treated DNA fragments from genes associated with DNA
 CC repair, and their complements.
 CC Note: The sequence data for this patent is not represented in the
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 13919 BP; 3717 A; 296 C; 3349 G; 6557 T; 0 other;
 Query Match 4.9%; Score 28; DB 24; Length 13919;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 547 TAATCTAAATAAAAAAAAAAAAAAAAAAAAAA 574
 Db 7862 TAATCTAAATAAAAAAAAAAAAAAAAAAAAAA 7835
 RESULT 14
 AAL30125/c
 ID AAL30125 standard; DNA; 51 BP.
 XX
 AC AAL30125;
 XX
 DT 24-JAN-2002 (first entry)
 DE Human SNP oligonucleotide #3333.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
 KW anyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.
 XX 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX (CURA-) CURAGEN CORP.
 PA Shimkets RA, Leach M;
 XX WPI; 2001-465210/50.
 DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 XX oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.
 PT Claim 1; Page 2342; 4143pp; English.
 PS
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, cancer
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 XX
 SQ Sequence 51 BP; 11 A; 1 C; 3 G; 36 T; 0 other;
 Query Match 4.7%; Score 27; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 548 AATCTAAATAAAAAAAAAAAAAAAAAAAAAA 574
 Db 32 AATCTAAATAAAAAAAAAAAAAAAAAAAAAA 6
 RESULT 15
 AAH33064
 ID AAH33064 standard; cDNA; 233 BP.
 XX
 AC AAH33064;
 XX
 DT 03-SEP-2001 (first entry)
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:120.
 XX
 DE Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; ss.
 KW
 KW Homo sapiens.
 OS
 XX WO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US26524.
 PF
 XX 29-SEP-1999; 99US-0157137.
 PR
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI

XX WPI: 2001-235357/24.
DR P-PSDB; AG73633.
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 2294; 9803pp; English.
XX AAH32943 to AAH37195 and AG73514 to AG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX SQ Sequence 233 BP; 116 A; 37 C; 40 G; 40 T; 0 other;

Query Match 4.7%; Score 27; DB 22; Length 233;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGCGTCGCCACCGCTCCGCA 27
Db 4 ACCACGCGTCGCCACCGCTCCGCA 30
|||||

Search completed: June 16, 2003, 06:54:05
Job time : 192.231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: June 16, 2003, 07:57:31 ; Search time 127.198 Seconds
(without alignments)
6534.570 Million cell updates/sec
Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 acccagcgctccgcccagcgc.....aaaaaaaaaaaaaaaaaaaaa 574
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 1029858 seqs, 724030393 residues
Word size: 0
Total number of hits satisfying chosen parameters: 2059716
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database: Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/PTCTUS_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	574	100.0	574	9	US-10-090-035-3
2	574	100.0	577	9	US-10-090-035-5
3	313	54.5	676	9	US-10-090-035-1
4	194	33.8	529	9	US-10-090-035-9
5	178	31.0	580	9	US-10-090-035-7
6	139	24.2	524	9	US-10-090-035-17
7	91	15.9	237	10	US-09-923-876-2788
8	44	7.7	348	9	US-10-090-035-13
9	44	7.7	591	9	US-10-090-035-15
10	34	5.9	436	9	US-10-090-035-21
11	34	5.9	584	9	US-10-090-035-19
12	31	5.4	908	10	US-10-090-035-23
13	29	5.1	538	9	US-09-800-729-25
c 14	29	5.1	569	10	US-09-918-995-17655
c 15	29	5.1	569	10	US-09-917-800A-508
16	28	4.9	405	9	US-09-918-995-16259
17	28	4.9	481	9	US-09-918-995-6561
18	28	4.9	532	9	US-09-776-724A-24
19	28	4.9	816	9	US-10-125-237-28

20	28	4.9	816	9	US-10-105-891-28
21	28	4.9	2674	9	US-10-098-841-148
22	28	4.9	4929	9	US-10-037-270-674
23	27	4.7	390	9	US-09-918-995-17194
24	27	4.7	393	9	US-09-918-995-7993
c 25	27	4.7	394	9	US-10-060-036-2561
26	27	4.7	400	9	US-09-925-299-465
27	27	4.7	400	10	US-09-925-299-265
28	27	4.7	421	9	US-09-918-995-16314
29	27	4.7	421	9	US-09-918-995-2984
30	27	4.7	451	9	US-10-072-349-76
31	27	4.7	451	10	US-09-764-855-76
32	27	4.7	468	9	US-09-918-995-5350
33	27	4.7	471	10	US-09-967-768A-69
34	27	4.7	471	10	US-09-864-761-5382
35	27	4.7	488	9	US-09-918-995-16810
36	27	4.7	508	9	US-09-918-995-16651
c 37	27	4.7	583	10	US-09-919-580-666
38	27	4.7	1263	10	US-09-785-770A-7
39	27	4.7	1794	10	US-09-925-301-426
40	27	4.7	1837	9	US-09-813-153-72
41	27	4.7	1927	9	US-10-138-846-11355
c 42	27	4.7	7802	9	US-10-239-676-56
43	27	4.7	10445	9	US-09-764-891-6380
44	27	4.7	16892	9	US-09-764-872-642
45	26	4.5	164	10	US-09-925-302-27

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090, 035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Query Match	100.0%	Score 574;	DB 9;	Length 574;
Best Local Similarity	100.0%	Pred. No. 2.5e-278;		
Matches 574;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACCCAGCGTCCGCCAGCGCTCCGCACAGCAATCCACAAAGCACTTCGACGTCACACG	60	
DB	1	ACCCAGCGTCCGCCAGCGCTCCGCACAGCAATCCACAAAGCACTTCGACGTCACACG	60	
QY	61	GGCGTGGCGCACAGACACACCAAGCGTCGGCAGCAATCCACAAAGCACTTCGACGTCACACG	120	
DB	61	GGCGTGGCGCACAGACACACCAAGCGTCGGCAGCAATCCACAAAGCACTTCGACGTCACACG	120	
QY	121	ACTGCTCGGAGAGTGAGTTCGGTGGCGCGCGCGCTTCGGCGCGCGCGCGCGCGCGCGG	180	
DB	121	ACTGCTCGGAGAGTGAGTTCGGTGGCGCGCGCGCTTCGGCGCGCGCGCGCGCGCGCGG	180	
QY	181	TCCAGCAGCAGTTCGTAAGGAGAGTTCGAGGAGTTCGACAGGTCGACAGGTCCTACGCGCGCGG	240	
DB	181	TCCAGCAGCAGTTCGTAAGGAGAGTTCGAGGAGTTCGACAGGTCGACAGGTCCTACGCGCGCGG	240	

db	383	CCGGCATCCACCCCGTTCGTGGCTGGTGGCCCTATGATGTCGTGGTGGACTGG	442
Qy	450	TTGTTTCAGGGTCATCGTACTTGGCTATCGTACGTGCACGCACTCAGCTCCTGTACGAAT	509
db	443	TTGTGCAGGGTCATCGTACTTGGCTATCGTACGTGCACGCACTCAGCTCCTGTACGAAT	502
Qy	510	ACGACATAAGTCGTGACCTGAATAAACCTTCTCGTAATACTAA	555
db	503	ACGACATAAGTCGTGACCTGAATAAACCTTCTCGTAATACTAA	548

RESULT 4

```

US-10-090-035-9
;
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(331)
US-10-090-035-9

```

Query Match	33.8%	Score 194;	DB 9;	Length 529;
Best Local Similarity	99.3%;	Pred. No. 1.2e-87;		
Matches 294;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	84	GCCTCGCGACCAATGGCTTACTACCAAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCG	143	
Db	41	GCGTGCGGACCAATGGCTTACTACCAAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCG	100	
QY	144	GTGGCCCGCGCGGCTTCGCGCGCCACACGCGCGGCGGCTCGACGACACGTCGTCAAGAG	203	
Db	101	GTGGCCCGCGCGGCTTCGCGCGCCACGCGCGGCGGCTCGACGACACGTCGTCAAGAG	160	
QY	204	AAGTTCGAGAGGTCGACACGGTCTCACGCGCGCGGCGCCACACGACACACACACGTCGT	263	
Db	161	AAGTTCGAGAGGTCGACACGGTCTCACGCGCGCGGCGCCACACGACACACACGTCGT	220	
QY	264	CACCACGGCGGCCACGGCTTCGTGTGCGGAGACACGAGGTGCGAAGAGGACATCAACACC	323	
Db	221	CACCACGGCGGCCACGGCTTCGTGTGCGGAGACACGAGGTGCGAAGAGGACATCAACACC	280	
QY	324	TGCACCGGCGAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTAACTGAGC	379	
Db	281	TGCACCGGCGAGGTCCACGAGCGGAGGAGAGCTTCCTCGCCAGGGCTAACTGAGC	336	

RESULT 5

: Sequence 7, Application US/10090035
 : Patent No. US20020170089A1
 : GENERAL INFORMATION:
 : APPLICANT: Simmons, Carl R.
 : TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
 : TITLE OF INVENTION: Proteins and Uses Thereof
 : FILE REFERENCE: 35718/242390
 : CURRENT APPLICATION NUMBER: US/10/090,035
 : CURRENT FILING DATE: 2002-02-28
 : PRIOR APPLICATION NUMBER: 60/272,227
 : PRIOR FILING DATE: 02/28/2001

```

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for W
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(380)
US-10-090-035-7

```

Query Match	31.0%	Score 178;	DB 9;	Length 580;
Best Local Similarity	99.1%	Pred. No. 1.3e-79;		
Matches 328;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	244	ACCACACACACACCATGGTCCACACGGCGGCCACCGCTTCGTGGTCCGACGACACGGG	303	
Db	250	ACCACACACACACCATGGTCCACACGGCGGCCACCGCTTCGTGGTCCGACGACACGGG	309	
QY	304	TCGAAGAGGACATCAACACCTGCACCGCGCGGAGGTCCACGAGCGCAGGGAGAGCTTCTCTCG	363	
Db	310	TCGAGGAGGACATCAACACCTGCACCGCGCGGAGGTCCACGAGCGCAGGGAGAGCTTCTCTCG	369	
QY	364	CGAGGCTAACTCAGCGCGCGCGCGCGCGATCCACGCCCGTTCGTGGTTCGCTCGGTG	423	
Db	370	CGAGGCTAACTCAGCGCGCGCGCGCGCGATCCACGCCCGTTCGTGGTTCGCTCGGTG	429	
QY	424	CCTTATGTATGTCGTGGTGTACTGGTGTTCAGGGTCATCGTACCTTGGCTATCGTACGT	483	
Db	430	CCTTATGTATGTCGTGGTGTACTGGTGTTCAGGGTCATCGTACCTTGGCTATCGTACGT	489	
QY	484	GCACGCACCTCAGCTCCTCTACGAATTACGACAATAAGCTCGTGACCTGAATAAACTTCT	543	
Db	490	GCACGCACCTCAGCTCCTCTACGAATTACGACAATAAGCTCGTGACCTGAATAAACTTCT	549	
QY	544	TCGTAATACTAAAAAATAAAAAAATAAAAAA	574	
Db	550	TCGTAATACTAAAAAATAAAAAAATAAAAAA	580	

RESULT 6

```

US-10-090-035-17
; Sequence 17, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/2772,277
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(338)
; NAME/KEY: misc_feature
; LOCATION: (1)...(524)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-17

```

Query Match 24.2%; Score 139; DB 9; Length 524;
Best Local Similarity 99.5%; Pred. No. 4.7e-60;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0

70 CACAGACACACCAAGCGTGGCGACCAATGGCTTACTTACCAGGAGGTGGACTACTCTCGG 129
|||||

```

Db 31 CACAGACACACCAAGCGTGGCCGACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 90
QY 130 AGGAGGTAGGTGCGTGGCCCGCGGCGTTCGGCCGCCACAGCGCGCGGCGTTCACGACG 189
Db 91 AGGAGGTAGGTGCGTGGCCCGCGGCGTTCGGCCGCCACAGCGCGCGGCGTTCACGACG 150
QY 190 AGTGGTCAAGAGAGTTCGAGGAGTTCGACACGCTCTCAGCGCGCGCGCCACACAC 249
Db 151 ACCTGCTCAAGAGAGTTCGAGGAGTTCGACACGCTCTCAGCGCGCGCGCCACACAC 210
QY 250 ACCACCA 259
Db 211 ACCACCA 220

```

```

RESULT 7
US-09-923-876-2788
; Sequence 2788, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2788
; LENGTH: 237
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: incyte ID No. US20020013958A1 700161180H1
; LOCATION: 125-146, 202
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-2788

```

```

Query Match 15.9%; Score 91; DB 10; Length 237;
Best Local Similarity 100.0%; Pred. No. 6.2e-36;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 CACAGACACACCAAGCGTGGCCGACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 129
Db 34 CACAGACACACCAAGCGTGGCCGACCAATGGCTTACTACAGGAGGTGGACTACTGCTCGG 93

QY 130 AGGAGGTAGGTGCGTGGCCCGCGGCGTTCGGCCGCCACAGCGCGCGGCGTTCACGACG 160
Db 94 AGGAGGTAGGTGCGTGGCCCGCGGCGTTCGGCCGCCACAGCGCGCGGCGTTCACGACG 124

```

```

RESULT 8
US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc.feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-13

```

```

Query Match 7.7%; Score 44; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACCAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCTCGGTG 146
Db 59 ACTACCAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCTCGGTG 102

```

```

RESULT 9
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc.feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15

```

```

Query Match 7.7%; Score 44; DB 9; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ACTACCAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCTCGGTG 146
Db 68 ACTACCAGGAGTGGACTACTGCTCGGAGGAGGTGAGTCTCGGTG 111

```

```

RESULT 10
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21

```

; LENGTH: 436
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(326)
US-10-090-035-21

Query Match 5.9%; Score 34; DB 9; Length 436;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 140
Db 65 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 98

RESULT 11

US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G

US-10-090-035-19

Query Match 5.9%; Score 34; DB 9; Length 584;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 140
Db 57 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 90

RESULT 12

US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-23

Query Match 5.9%; Score 34; DB 9; Length 584;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 140
Db 57 CCAGGAGGTGACTCTCGGAGGAGGTGAGG 90

RESULT 13

US-09-800-729-25
; Sequence 25, Application US/09800729.
; Patent No. US20020068319A1
; GENERAL INFORMATION:

; APPLICANT: NI et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (891)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (896)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-800-729-25

Query Match 5.4%; Score 31; DB 10; Length 908;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAGCGGTCCGCCACGCGTCCGCACAGC 31
Db 19 ACCCAGCGGTCCGCCACGCGTCCGCACAGC 49

RESULT 14

US-09-918-995-17655/c

; Sequence 17655, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17655

LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(538)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17655

Query Match 5.1%; Score 29; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAGCGTCGCCCGCAGCGTCGCCGACA 29
Db 209 ACCCAGCGTCGCCCGCAGCGTCGCCGACA 181

RESULT 15

US-09-917-800A-508/c
Sequence 508, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castile, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 508
LENGTH: 569
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 AT007824
US-09-917-800A-508

Query Match 5.1%; Score 29; DB 10; Length 569;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 GTATACTAAAAAATAAAAAAAAAA 574
Db 73 GTATACTAAAAAATAAAAAAAAAA 45

Search completed: June 16, 2003, 10:24:26
Job time : 129.198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 05:14:00 ; Search time 1370.39 Seconds
(Without alignments)
7989.092 Million cell updates/sec

Title: US-10-090-035-1

Perfect score: 676

Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	308	45.6	500	10	AW787732
2	305	45.1	524	10	AW288876
3	279	41.3	465	10	BE129897
4	268	39.6	523	10	AW331212
5	261	38.6	553	11	AV104409
6	261	38.6	648	14	BQ619167

7	261	38.6	648	14	BQ619315
8	261	38.6	648	14	BQ619318
9	261	38.6	648	14	BQ619337
10	261	38.6	648	14	BQ619383
11	261	38.6	648	14	BQ619390
12	256	37.9	550	12	BG840383
13	255	37.7	546	9	AA979839
14	237	35.1	417	13	BM501439
15	232	34.3	436	10	AW787315
16	226	33.4	552	10	BE025302
17	214	31.7	466	9	AI964534
18	214	31.7	496	10	AW787314
19	200	29.6	539	9	AI855425
20	198	29.3	481	9	AI964458
21	197	29.1	443	10	BE129644
22	188	27.8	401	10	AW289056
23	168	24.9	476	10	BE519299
24	152	22.5	326	10	BE025303
25	137	20.3	274	10	AW288875
26	122	18.0	173	14	BM885464
27	122	18.0	215	13	BM530859
28	122	18.0	220	13	BM154814
29	122	18.0	222	14	BQ085944
30	122	18.0	244	13	BM154314
31	122	18.0	1011	14	BQ954179
32	122	18.0	1030	14	BQ222049
33	121	17.9	240	13	BJ389253
34	121	17.9	242	10	AW781584
35	121	17.9	242	10	BE059705
36	121	17.9	275	12	BE777697
37	121	17.9	305	12	BE777697
38	121	17.9	384	9	AL703622
39	121	17.9	399	9	AI226246
40	121	17.9	402	12	BE726297
41	121	17.9	416	10	AW885721
42	121	17.9	711	12	BG109125
43	121	17.9	768	10	AV755678
44	121	17.9	806	10	AV757705
45	121	17.9	842	10	AV757327

ALIGNMENTS

RESULT 1
AW787732

LOCUS
DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AW787732

945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707

(SK) Zea mays cDNA, mRNA sequence.

AW787732

AW787732.1 GI:7844510

EST.

Zea mays.

Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

1 (bases 1 to 500)

Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 945002 row: E column: 06.

Location/Qualifiers

1..500

/organism="Zea mays"

500 bp mRNA linear EST 16-MAY-2000

...

clade: Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 465)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1. 465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site: 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

88 a 151 c 149 g 77 t

BASE COUNT
ORIGIN

Query Match 41.3%; Score 279; DB 10; Length 465;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
DB 34 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 93

QY 143 CCGCGCGGCTTCGGCGCCGACGGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 202
DB 94 CCGCGCGGCTTCGGCGCCGACGGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 153

QY 203 GAGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 262
DB 154 GAGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 213

QY 263 GCGCGCCACGGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 322
DB 214 GCGCGCCACGGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 273

QY 323 GCGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 382
DB 274 GCGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 333

QY 383 CCGCATCCACGCGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 442
DB 334 CCGCATCCACGCGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 393

QY 443 TTGTGAGGATTCATCGTACTTGGCTATCGTACGTCGACGCTCAAGGAGATTC 502
DB 394 TTGTGAGGATTCATCGTACTTGGCTATCGTACGTCGACGCTCAAGGAGATTC 502

QY 503 ACCACAAATAGC 514
DB 454 ACCACAAATAGC 465

RESULT 4
AW331212
LOCUS

523 bp mRNA linear EST 31-JAN-2000

DEFINITION
1 (bases 1 to 465)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1. 465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site: 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

88 a 151 c 149 g 77 t

BASE COUNT
ORIGIN

Query Match 39.6%; Score 268; DB 10; Length 523;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 468; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 83 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 142
DB 29 GCACCAATGGCTTACTACCAAGAGGTGACTGCTCGGAGGAGGTGAGTGGTGGCC 88

QY 143 CCGCGCGGCTTCGGCGCCGACGGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 202
DB 89 CCGCGCGGCTTCGGCGCCGACGGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 148

QY 203 GAGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 262
DB 149 GAGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 208

QY 263 GCGCGCCACGGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 322
DB 209 GCGCGCCACGGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 268

QY 323 GCGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 382
DB 269 GCGGAGTTCGACAGGTATCACGCGCGCGGCCAACCCACCACCATGCTCAAGGAGATTC 328

QY 383 CCGCATCCACGCGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 442
DB 329 CCGCATCCACGCGCTTCGTTGGTTCGCGAGACCGAGGTTCCTCGCCAGGCTTAACGAGCGCGCGGCC 388

QY 443 TTGTGAGGATTCATCGTACTTGGCTATCGTACGTCGACGCTCAAGGAGATTC 502
DB 389 TTGTGAGGATTCATCGTACTTGGCTATCGTACGTCGACGCTCAAGGAGATTC 448

FEATURES	source
Location/Qualifiers	
1. .436	
/organism="Zea mays"	
/cultivar="W23"	
/db_xref="taxon:4577"	

TITLE	Maize opaque endosperm mutations create extensive changes in patterns of gene expression
JOURNAL	Unpublished (2002)
COMMENT	Contact: Jung R Trait and Technology Development, Food and Feed Research Pioneer Hi-Bred International, Inc.

```

/clone.lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue.type="tassel, kernal, silk, husk, root, leaf"
/dev.stage="fully-grown"
/lab.host="DHI0B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT      82 a 136 c 136 g 82 t
ORIGIN

```

```

Query Match      34.3%; Score 232; DB 10; Length 436;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 432; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 108 TGGACTACTGTCGGAGGAGTGAGTGGCGCCCGCGGCTTCGGCCGCCACGGCG 167
Db 1 TGGACTACTGTCGGAGGAGTGAGTGGCGCCCGCGGCTTCGGCCGCCACGGCG 60

Qy 168 GCGCGCTCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGGC 227
Db 61 GCGCGCTCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTCTCACGGC 120

Qy 228 CCGCGCCCAACCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTATCACGGC 227
Db 121 CCGCGCCCAACCCAGCAGCAGTCTCAAGAGAGAGTTTCGAGAGGTCGACACGGTCTCACGGC 180

Qy 288 AGACCCAGGTCGAGGAGGAGATCAACACCTGCACCGCGGAGTTCACAGCGCAGGAGAG 347
Db 181 AGACCCAGGTCGAGGAGGAGATCAACACCTGCACCGCGGAGTTCACAGCGCAGGAGAG 240

Qy 348 GCTTCCTCGCCAGGGCTAACTAGCGCGCGCGCGCGGATCCACGCCCGCTTCGTGCTT 407
Db 241 GCTTCCTCGCCAGGGCTAACTAGCGCGCGCGCGGATCCACGCCCGCTTCGTGCTT 300

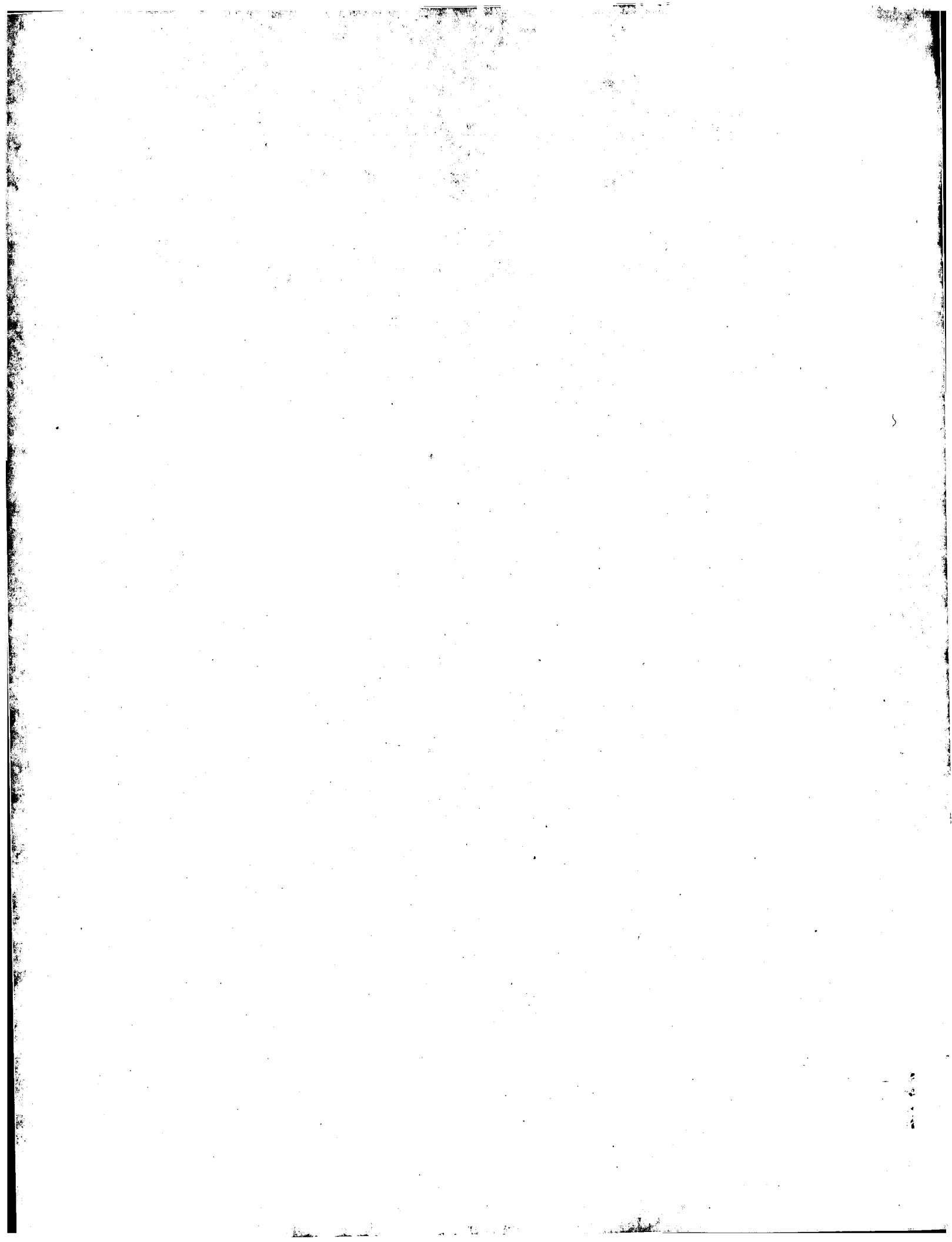
Qy 408 GCCTGGCTGCTTATGTATGTCTGTGTTGACTGTTGTGAGGGTTCATCGTACTTGGCT 467
Db 301 GCCTGGCTGCTTATGTATGTCTGTGTTGACTGTTGTGAGGGTTCATCGTACTTGGCT 360

Qy 468 ATCGTACGTGCACGCCACTCAGCTCCTGTACGATTAACGACAATAGCTCGTGACCTGAAT 527
Db 361 ATCGTACGTGCACGCCACTCAGCTCCTGTACGATTAACGACAATAGCTCGTGACCTGAAT 420

Qy 528 AAAAATTCCTCGTAAT 543
Db 421 AAAAATTCCTCGTAAT 436

```

Search completed: June 16, 2003, 08:38:41
Job time : 1384.39 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 10:02:07 ; Search time 1293.5 Seconds
(without alignments)
1164.422 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510

Sequence: 1 MAYQEVDCYSEVRVAPA.....INTCTGVHERRESFLARAN 93

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+_p2n.model -DEV-xlh
-Q/cgn2_1/USPTO_spool/US10090035/runat_06062003_105504_10970/app_query.fasta_1.526
-DB-EST -QFMT-fastap -SUFFIX-rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -NATRIX-blosum62 -TRANS-human40.cdi -LIST-45
-DOALIGN-200 -THR SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HPAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US10090035 -CGN_1_2463 -runat_06062003_105504_10970 -NCPU-6 -ICPU-3
-NO_WMAP -LARGQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

EST:*
1: em_estbn:*
2: em_esthm:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	510	100.0	417	13	BM501439	BM501439 PAC000000
2	505	99.0	465	10	BE129897	BE129897 945032C12
3	505	99.0	500	10	AW787732	AW787732 945002E06
4	505	99.0	523	10	AW331212	AW331212 707049E04
5	505	99.0	524	10	AW288876	AW288876 707009E07
6	499.5	97.9	466	9	AI964534	AI964534 496013D01
C 7	499.5	97.9	481	9	AI964458	AI964458 496010H04
C 8	499.5	97.9	539	9	AI855425	AI855425 603016F02
9	499.5	97.9	546	9	AA979839	AA979839 MEST2-B7
10	499.5	97.9	550	12	BG840383	BG840383 MEST12-H1
11	499.5	97.9	553	11	AY104409	AY104409 Zeb_mays
12	499.5	97.9	648	14	BQ619167	BQ619167 RNOSEQ4E0
13	499.5	97.9	648	14	BQ619135	BQ619135 RNOSEQ6C0
14	499.5	97.9	648	14	BQ619318	BQ619318 RNOSEQ6D0
15	499.5	97.9	648	14	BQ619337	BQ619337 RNOSEQ6E1
16	499.5	97.9	648	14	BQ619383	BQ619383 RNOSEQ7B0
17	499.5	97.9	648	14	BQ619390	BQ619390 RNOSEQ7B0
18	498	97.6	326	18	BE025303	BE025303 945028B09
19	496	97.3	443	10	BE129644	BE129644 945027E06
20	496	97.3	552	10	BE025302	BE025302 945028B09
21	494	96.9	476	10	BE519299	BE519299 945007B05
22	492	96.5	496	10	AW787314	AW787314 945002E06
23	482	94.5	433	10	BE225008	BE225008 945042F02
24	468	91.8	436	10	AW787315	AW787315 945002E06
25	423.5	83.0	399	13	BM318672	BM318672 P11_16_C0
26	423.5	83.0	466	10	BE597738	BE597738 P11_85_F0
27	423.5	83.0	485	10	BE599123	BE599123 P11_85_F0
28	423.5	83.0	513	14	BQ280709	BQ280709 WHE3004_A
29	423.5	83.0	516	10	BE364814	BE364814 P11_16_C0
30	423.5	83.0	519	10	AW680016	AW680016 WSL_34_H1
31	423.5	83.0	522	10	AW679969	AW679969 WSL_34_H1
32	423.5	83.0	537	10	AW745400	AW745400 WSL_34_H1
33	423.5	83.0	541	10	AW745436	AW745436 WSL_34_B0
34	417.5	81.9	348	10	BE593507	BE593507 WSL_100_B
35	410.5	80.5	509	14	BQ280894	BQ280894 WHE3006_B
36	395	77.5	401	10	AW283056	AW283056 707005E07
37	368.5	72.3	274	10	AW288875	AW288875 707009E07
38	367.5	72.1	311	12	BF729420	BF729420 100007C0
39	308	60.4	453	10	AW923922	AW923922 WSL_30_A1
40	308	60.4	554	10	AW679915	AW679915 WSL_33_G0
41	308	60.4	560	10	AW924079	AW924079 WSL_49_G0
42	308	60.4	566	10	AW677917	AW677917 WSL_12_D0
43	308	60.4	566	10	AW746383	AW746383 WSL_49_G0
44	308	60.4	572	10	AW679713	AW679713 WSL_30_A1
45	308	60.4	572	10	AW925014	AW925014 WSL_74_D0

ALIGNMENTS

RESULT 1
BM501439

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BM501439 417 bp mRNA linear EST 14-FEB-2002
PAC000000000593 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.

BM501439

BM501439.1 GI:18661517

EST

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 417)

Hunter, B.G., Beatty, M., Singletary, G., Hamaker, B., Larkins, B.A. and

Jung, R.

TITLE Maize opaque endosperm mutations create extensive changes in patterns of gene expression
JOURNAL Unpublished (2002)
COMMENT Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.

FEATURES
source Location/Qualifiers
1. .417
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI"

BASE COUNT 81 a 150 c 133 g 53 t
ORIGIN

Alignment Scores:
Pred. No.: 5.6e-47 Length: 417
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-090-035-4 (1-93) x BM501439 (1-417)

QY 1 MetAlaTYrTYrGlnGluValAspTYrCysSerGluGluValArgSerValAlaProAla 20
Db 85 ATGGCTTACTACGAGGAGTGGACTACTGCTCGAGGAGGTGAGTGGTGGCCCGGCC 144
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 145 GGCTTGGCGCGCCACGGCGCGCGCTCCAGCACGTGCTCAAGGAGAGTTCGAGGAG 204
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyGly 60
Db 205 GTTCACACGGTCTCACGGCGCGCGCCACACCATGTGTCCACACGGCGC 264
QY 61 HisGlyPheValValArgGluThrArgValGluAspIleAsnThrCysThrGlyGlu 80
Db 265 CACGGCTTGTGTGTCGCGAGACCGAGGTGCGAGAGGACATCAACACCTGCACCGCGAG 324
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 325 GTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 363

RESULT 2
BE129897 465 bp mRNA linear EST 21-JUN-2000
LOCUS 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE129897.1 GI:8577260
VERSION BE129897
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
JOURNAL clade; Panicoideae; Andropogoneae; Zea.
COMMENT 1 (bases 1 to 465)
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

Email: walbot@stanford.edu
Plate: 945032 row: C column: 12.
Location/Qualifiers
1. .465
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="Phi10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 88-a 151 c 149 g 77 t
ORIGIN

Alignment Scores:
Pred. No.: 2.26e-46 Length: 465
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-4 (1-93) x BE129897 (1-465)

QY 1 MetAlaTYrTYrGlnGluValAspTYrCysSerGluGluValArgSerValAlaProAla 20
Db 40 ATGGCTTACTACGAGGAGTGGACTACTGCTCGAGGAGGTGAGTGGTGGCCCGGCC 99
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 100 GCCTTGGCGCGCCACGGCGCGCGCTCCAGCACGTGCTCAAGGAGAGTTCGAGGAG 159
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyGly 60
Db 160 GTCCACACGGTCTCACGGCGCGCGCCACCATGTGTCCACACGGCGC 219
QY 61 HisGlyPheValValArgGluThrArgValGluAspIleAsnThrCysThrGlyGlu 80
Db 220 CACGGCTTGTGTGTCGCGAGACCGAGGTGCGAGGAGACATCAACACCTGCACCGCGAG 279
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 280 TTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 318

RESULT 3
AW787732 500 bp mRNA linear EST 16-MAY-2000
LOCUS 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787732
VERSION AW787732.1 GI:7844510
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
JOURNAL clade; Panicoideae; Andropogoneae; Zea.
COMMENT 1 (bases 1 to 500)
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E col:

FEATURES

```

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)."
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DHI08"
/notes="Organ: tassel, kernal, silk, husk, root, leaf;
vector: pGAD10; Site1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
97 a 159 g 88 t

```

BASE COUNT
ORIGIN

Alignment Scores:

2.45e-46	Length:	500
Score:	Matches:	92
Percent Similarity:	Conservative:	0
Best Local Similarity:	Mismatches:	1
Query Match:	Indels:	0
DB:	Gaps:	0

JS-10-090-035-4 (1-93) x AW787732 (1-500)

1	MetAlaTyrTyrGlnGlnValAlaSpTyrCysSerGluGluValArgSerValAlaProAla	20
35	ATGGCTTACTACCGAGGAGTGGACTACTCTCGGAGGAGGTGAGTCTGGTGGCCCCGGCC	94
21	GlyPheGlyArgHisGlyGlyValGlnGlnHisValLysGluLysPheGluGlu	40
95	GGCTTCGGCGCCACACGGCGCGCGGTCCAGCAGCACGTCGTCAAGGAGAAATTCTGAGGAG	154
41	ValAspThrValSerArgAlaGlyAlaAsnHisHisHisGlyHisHisGlyGly	60
135	GTCCGACACGGTCTACGCGCGCGCCACACCAACCACCATGTGTACACGCGGCG	214
61	HisGlyPheValValArgGluThrArgValGluAspIleAsnThrCysThrGlyGlu	80
215	CACGGCTTCGTGGTGGCGGACACAGGTCTGAAAGAGGACATCAACCTCACCGCGGAG	274
81	ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn	93
275	TTCCACGAGCGCAGGAGAGCTTCCTGCCACGGCGTAACT	313

RESULT 4
W331212

OCUS AW331212

DEFINITION
 707049804.xl 707 - Mixed adult tissues from Walbot lab (SK) Zea
 mays cDNA, mRNA sequence.
 ACCESSION AW331212
 VERSION AW331212.1 GI:6827569
 KEYWORDS EST.

SOURCE Zea mays.

ORGANISM

zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; zea.

1 (bases 1 to 523)

REFERENCE

AUTHORS	TITLE
Walbot, V.	Maize ESTs from various sources

**JOURNAL
COMMENT**

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707049 row: E column: 04.

FEATURES

```

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="DH108"
/lab_host="DH108"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site.1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

```

BASE COUNT	125 a	158 c	149 g	91 t	Unidirectional
muscle, foot, leaf					

Alignment Scores:

Pred. No.:	2,58-45	Length:	523
Score:	505.00	Matches:	92
Percent Similarity:	98.9%	Conservative:	0
Best Local Similarity:	98.9%	Mismatches:	1
Query Match:	99.0%	Indels:	0
DB:	10	Gaps:	0

US-10-090-035-4 (1-93) x AW331212 (1-523)

1	Meta	Ty	Ty	Ty	Gln	Glu	Val	Ser	Cys	Ser	Glu	Glu	Val	Arg	Ser	Val	Ala	Pro	Ala	20
35	AT	GG	CT	TACT	AC	AG	AG	GT	GG	ACT	TCT	GC	GG	AG	GT	CAG	GT	CG	TGG	CC
21	Gly	Ph	Gly	Arg	His	Gly	Gly	Val	Gln	Gln	His	Val	Val	Lys	Glu	Lys	Ph	Glu	Glu	40
95	GG	CT	CG	CG	CC	AG	CG	GG	CG	CT	CC	AG	AC	AG	CT	CG	TCA	MG	AG	AA
41	Val	Asp	Thr	Val	Ser	Arg	Ala	Gly	Ala	Asn	His	His	His	His	Gly	His	His	Gly	Gly	60
155	GT	CG	AC	GG	CT	CT	CAC	GC	CG	CG	CG	CA	ACC	ACC	ACC	ACC	ACC	ACC	ACC	
61	His	Gly	Ph	Val	Val	Arg	Glu	Thr	Arg	Val	Glu	Glu	Asp	Phe	Leu	Ala	Asn	Thr	Cys	
215	CAC	GC	CT	TC	GT	GT	GC	CG	AG	AC	CAG	GG	TC	GA	AG	AG	AC	AT	CA	
81	Val	His	Glu	Arg	Arg	Glu	Ser	Ph	Leu	Ala	Arg	Ala	Asn							
275	TT	CC	AC	AG	CG	AC	GG	AG	AG	CT	TCT	CG	CG	AC	GG	CG	CT	AA	CC	

RESULT 5

W288876

	OCUS	AW288876	524 bp	mRNA	linear	EST 16-JAN-2000
	EFINITION	707009E07.x4	707	Mixed adult tissues from Walbot lab (SK) zea mays cDNA, mRNA sequence.		

AW288876 1 CT.6605663

KEYWORDS
SOURCE
ORGANISM

2000

REFERENCE

AUTHORS

ATTI

JOURNAL

COMMENT

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707009 row: E column: 07.

FEATURES

Location/Qualifiers
1..524
/organism="Zea mays"
/cultivar="W23"

/db_xref="taxon:4577"
/clone_lib="707" - Mixed adult tissues from walbot lab (SK
)

/tissue_type="tassel, kernel, silk, husk, root, leaf"

/dev_stage="adult"

/lab_host="PH10B"

/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site: EcoRI; cDNA library from fully

differentiated maize tissues from an active Mutator

plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,

husk, root, leaf). Unidirectionally cloned."

BASE COUNT 107 a 170 c 157 g 89 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 2.58e-46 Length: 524
Score: 505.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 1
Query Match: 99.02% Indels: 0
DB: 10 Gaps: 0

US-10-090-035-4 (1-93) x AW288876 (1-524)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 73 ATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGCCCGCC 132
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 133 GGCATTCGGCGCCACGCGCGCGGTCCAGCAGCACGTCGTCGAAGAGAGTTCGAGGAG 192
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisGly 60
Db 193 GTCGACAGGCTCTACGCGCGCGCGCCCAACACCACCACCATGTCACCGCGCGC 252
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
Db 253 CACGGCTTCGTGTGTCGCGAGACACGAGGTGCAAGAGGACATCAACCTCGACCGCGGAG 312
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 313 TTCCACGAGCGAGGAGAGGCTTCTCGCCAGGGCTAAC 351

RESULT 6

AI964534/c 466 bp mRNA linear EST 20-AUG-1999
LOCUS 496013D01.x1 496 - stressed shoot cDNA library from Wang/Bohnert
DEFINITION lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964534
VERSION AI964534.1 GI:5757247
KEYWORDS EST.

SOURCE

ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 466)

Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496013 row: D column: 01.

FEATURES

Location/Qualifiers
1..466
/organism="Zea mays"
/cultivar="B73"

/db_xref="taxon:4577"
/clone_lib="496" - stressed shoot cDNA library from
Wang/Bohnert lab

/tissue_type="seedling"

/dev_stage="salt stress"

/lab_host="E.coli XL Gold"

/note="Organ: shoot; Vector: pBluescriptII SK(+) XR;

Wang/Bohnert"

BASE COUNT 76 a 151 c 154 g 85 t

ORIGIN

Alignment Scores:
Pred. No.: 9.2e-46 Length: 466
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AI964534 (1-466)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 416 ATGGCTTACTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGCCCGCC 357
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 356 GGCATTCGGCGCCACGCGCGCGGTCCAGCAGCACGTCGTCGAAGAGAGTTCGAGGAG 297
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisHisHisGly 59
Db 296 GTCGACAGGCTCTACGCGCGCGCCCAACACCACCACCATGTCACCGCGC 237
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 236 GGCACCGGCTTCGTGTGTCGCGAGACACGAGGTGCGAGGAGGACATCAACACCTGCACCGC 177

RESULT 7

AI964458/c 481 bp mRNA linear EST 20-AUG-1999
LOCUS 496010H04.x1 496 - stressed shoot cDNA library from Wang/Bohnert
DEFINITION lab Zea mays cDNA, mRNA sequence.

ACCESSION AI964458
VERSION AI964458.1 GI:5757171
KEYWORDS EST.

SOURCE

ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 481)

Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 496010 row: H column: 04.

FEATURES

source
1. .481
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="496 - stressed shoot cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: shoot; Vector: pBluescriptII SK(+/-) XR; Wang/Bohnert"

BASE COUNT 82 a 148 c 154 g 97 t
ORIGIN

Alignment Scores:
Pred. No.: 9 53e-46 Length: 481
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AI964458 (1-481)

QY 1 MetAlaTyrTyGInGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 451 ATGGCTTACTACACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTGGCCCGCC 392
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 391 GGCTTCGGCCCGCCAGGAGGCGGCTCCAGCACGCTCGTCAAGGAGAGTTCGAGGAG 332
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGlyHisGly 59
|||||
Db 331 GTCGACACGGTCTCACGCGCGCGCCGACCAACCACCACCACCATGTGTACACCGGC 272
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||
Db 271 GGCACCGCTCTGTGGTGGCGAGACACGAGGTCGAGGAGACATCAACACCTGCACCGGC 212
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 211 GAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 170

RESULT 8
AI855425/c
LOCUS
DEFINITION
603016F02.xl 603 - stressed root cDNA library from Wang/Bohnert lab
Zea mays cDNA, mRNA sequence.
ACCESSION
AI855425
VERSION
AI855425.1 GI:5499558
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 539)
Walbot.v.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.
Location/Qualifiers
1. .539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+/-) XR; Seedling stressed root cDNA library from Wang/Bohnert lab"

FEATURES

source
1. .539
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="Organ: root; Vector: pBluescriptII SK(+/-) XR; Seedling stressed root cDNA library from Wang/Bohnert lab"

BASE COUNT 91 a 157 c 172 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 1 08e-45 Length: 539
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AI855425 (1-539)

QY 1 MetAlaTyrTyGInGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 477 ATGGCTTACTACACGAGGAGTGACTACTGCTCGGAGGAGTGAGTGGTGGCCCGCC 418
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
|||||
Db 417 GGCTTCGGCCCGCCAGGAGGCGGCTCCAGCACGCTCGTCAAGGAGAGTTCGAGGAG 358
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGlyHisGly 59
|||||
Db 357 GTCGACACGGTCTCACGCGCGCGCCGACCAACCACCACCACCATGTGTACACCGGC 298
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
|||||
Db 297 GGCACCGCTCTGTGGTGGCGAGACACGAGGTCGAGGAGACATCAACACCTGCACCGGC 238
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
|||||
Db 237 GAGGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 196

RESULT 9

AA979839
LOCUS
DEFINITION
MEST2-B7.TW1412.Seq ISUM2 Zea mays cDNA clone MEST2-B7 5', mRNA
sequence.
AA979839
ACCESSION
AA979839.1 GI:3157217
VERSION
AA979839.1
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 546)
Wen,T.J., Ashlock,D.A. and Schnable,P.S.
Expressed Sequence Tags from B73 Maize Seedlings
Unpublished (1997)
Contact: Schnable, PS
Schnable laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975
Fax: (515)-294-2299
Email: schnable@iastate.edu

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
Expressed Sequence Tags from B73 Maize Seedlings
Unpublished (1997)
Contact: Schnable, PS
Schnable laboratory
Iowa State University
G405 Agronomy, Ames, IA 50011, USA
Tel: (515)-294-0975
Fax: (515)-294-2299
Email: schnable@iastate.edu

Mon Jun 16 14:55:59 2003

AUTHORS TITLE JOURNAL COMMENT

Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206705.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

FEATURES

source
1..546
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST2-B7"
/clone_lib="ISUM2"
/tissue_type="above ground tissues"
/dev_stage="Two-leaf-stage green seedling"
/lab_host="XLA-MFR Blue"
/note="Organ: green seedlings; Vector: pAD-GAL4; Site_1:
ECORI; Site_2: XhoI; ds-cDNA molecules were generated as
follows. First-strand cDNA was prepared from oligo-dT
selected mRNA by priming with an XhoI oligo-dT primer. The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with XhoI and size-selected. The
resulting molecules were directionally cloned into the
ECORI and XhoI sites of the HybridZAP lambda vector
(Stratagene) and excised as pAD-GAL4 phagemids."

BASE COUNT 130 a 168 c 159 g 88 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1..le-45 Length: 546
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 9 Gaps: 1

US-10-090-035-4 (1-93) x AA979839 (1-546)

Qy 1 MetaLatyrTyrgInGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 66 ATGGCTTACTACACAGAGGTGGACTACTGCTCGGAGGAGTGAGGTGCGGCCCGGCC 125
Qy 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 126 GGCTTCGGCGCCACGAGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAG 185
Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
Db 186 GTCGACACGGTCTCACGCGCGCGCCACACACACACACCATGTGTACACGCGC 245
Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 246 GCCCAGCGCTTCGTGGTGGCGGAGCACCAGGTCGAGGAGGACATCAACACCTGCACCGGC 305
Qy 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 306 GAGGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 347

RESULT 10
BG840383 550 bp mRNA linear EST 29-MAY-2001
LOCUS MEST12-H11.T7-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA
DEFINITION
ACCESSION BG840383
VERSION BG840383.2 GI:14242676
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 550)

PCR Primers FORWARD: T7-1 (5'-GAAGATACCCCAACCAACC-3') BACKWARD: T7-YJ (5'-TAATACGACTCATATAGGC-3') Plate: MEST2 row: B column: 7 Seq primer: tw1412 (5'-GAAGATACCCCAACCAACC-3').

Location/Qualifiers
1..550
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST12-H11"
/clone_lib="ISUM4-TN"
/tissue_type="Seedling and silk"
/lab_host="DH10B"
/note="vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'-
AACTGGAAGATTCCGGCGCGAGGAATTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT7T3PAC vector. The library
then went through one round of normalization to Cot value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

FEATURES

source

1..550

/organism="Zea mays"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="MEST12-H11"

/clone_lib="ISUM4-TN"

/tissue_type="Seedling and silk"

/lab_host="DH10B"

/note="vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;
ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'-
AACTGGAAGATTCCGGCGCGAGGAATTTTTTTTTTTTTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT7T3PAC vector. The library
then went through one round of normalization to Cot value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

BASE COUNT 124 a 174 c 159 g 92 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1..le-45 Length: 550
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 12 Gaps: 1

US-10-090-035-4 (1-93) x BG840383 (1-550)

Qy 1 MetaLatyrTyrgInGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 75 ATGGCTTACTACACAGAGGTGGACTACTGCTCGGAGGAGTGAGGTGCGGCCCGGCC 134
Qy 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 135 GGCTTCGGCGCCACGAGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGAG 194
Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisGly 59
Db 195 GTCGACACGGTCTCACGCGCGCGCCACACACACACCATGTGTACACGCGC 254
Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 255 GCCCAGCGCTTCGTGGTGGCGGAGCACCAGGTCGAGGAGGACATCAACACCTGCACCGGC 314
Qy 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 315 GAGGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGGCTAAC 356

RESULT 11
AY104409

LOCUS AY104409 553 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PC0124784 mRNA sequence.
 ACCESSION AY104409
 VERSION AY104409.1 GI:21207487
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays

REFERENCE 1 (bases 1 to 553)
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 553)
 AUTHORS Coe,E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

FEATURES
 source
 1..553
 Location/Qualifiers
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone_lib="PC0124784"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 128 a 173 c 162 g 90 t

Alignment Scores:
 Pred. No.: 1.11e-45 Length: 553
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 11 Gaps: 1

US-10-090-035-4 (1-93) x AY104409 (1-553)

Qy 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 76 ATGGCTTACTACCGAGGAGTGGACTCTCTCGGAGGAGTGGAGTGGTGGCCCGGCC 135
 Qy 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 136 GGCTTCGGCCGCCAGCGAGGCGGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 195
 Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisHisGly 59
 |||||
 Db 196 GTGCACACGGTCTCAGCGCCGCCAGCGGCGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 255
 Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
 |||||
 Db 256 GGCCACGGTCTCTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 315
 Qy 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 316 GAGGTCCACGAGCGCGAGGAGAGCTTCCTCGCCAGGGCTAAC 357

RESULT 12

LOCUS BQ619167 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ4E05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RN05EQ4E05_SK.ab1 similar to No homology, mRNA

sequence.
 BQ619167
 VERSION BQ619167.1 GI:21621161
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

REFERENCE 1 (bases 1 to 648)
 AUTHORS Wang,H. and Bohner,H.J.
 TITLE Genomics of plant stress tolerance
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mark Fredricksen
 Department of Plant Biology
 University of Illinois
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 2172653473
 Email: bohnerlab@life.uiuc.edu.

FEATURES
 source
 1..648
 Location/Qualifiers
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone_lib="RN05EQ4E05_SK.ab1"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /tissue_type="Roots"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

Alignment Scores:
 Pred. No.: 1.33e-45 Length: 648
 Score: 499.50 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 0
 Query Match: 97.94% Indels: 1
 DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x BQ619167 (1-648)

Qy 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
 |||||
 Db 82 ATGGCTTACTACCGAGGAGTGGACTCTCTCGGAGGAGTGGAGTGGTGGCCCGGCC 141
 Qy 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
 |||||
 Db 142 GGCTTCGGCCGCCAGCGAGGCGGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 201
 Qy 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisGlyHisHisGly 59
 |||||
 Db 202 GTGCACACGGTCTCAGCGCCGCCAGCGGCGCTCCAGCAGCAGCTCGTCAAGGAGAGTTCGAGGAG 261
 Qy 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
 |||||
 Db 252 GGCCACGGTCTCTGGTGGCGGAGACCGAGGTCGAGGAGGACATCAACACCTGCACCGGC 321
 Qy 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
 |||||
 Db 322 GAGGTCCACGAGCGCGAGGAGAGCTTCCTCGCCAGGGCTAAC 363

RESULT 13

LOCUS BQ619315 648 bp mRNA linear EST 27-JUN-2002
 DEFINITION RN05EQ6C06_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RN05EQ6C06_SK.ab1 similar to No homology, mRNA

sequence.
 BQ619315
 VERSION BQ619315.1 GI:21621309
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays

Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

FEATURES
source
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/dev_stage="2 weeks old"
/tissue_type="Roots"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x B0619318 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAlaSerGluGluValArgSerValAlaProAla 20
Db 82 ATGCTTACTACAGAGGTGGTCTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 142 GCTTCGGCCGCGGAGGCGGTCGACGAGCAGTCTCAAGGAGAGTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisGly 59
Db 202 GTGACACGCTCTACGCGCGCGCCCAACACCACCACCATGTCACACGCGC 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCTGTGTGTCGCGAGACAGGTCGAGGAGGAGCATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGGAGGAGGAGTCTCTCGCCAGGCGCTAAC 363

RESULT 15
B0619337
LOCUS
DEFINITION
mays cDNA clone RNOSEQ6E12_SK.ab1 similar to No homology, mRNA
sequence.

ACCESSION B0619337
VERSION B0619337.1 GI:21621331
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang, H. and Bohnert, H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

FEATURES
source
1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/dev_stage="2 weeks old"
/tissue_type="Roots"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

BASE COUNT 172 a 193 c 183 g 100 t

ORIGIN

Alignment Scores:
Pred. No.: 1.33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x B0619315 (1-648)

QY 1 MetAlaTyrTyrGlnGluValAlaSerGluGluValArgSerValAlaProAla 20
Db 82 ATGCTTACTACAGAGGTGGTCTCTCGGAGGAGGTGAGTGGTGGCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValysGluLysPheGluGlu 40
Db 142 GCTTCGGCCGCGGAGGCGGTCGACGAGCAGTCTCAAGGAGAGTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisHisGly 59
Db 202 GTGACACGCTCTACGCGCGCGCCCAACACCACCACCATGTCACACGCGC 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GCCCAGCGCTTCTGTGTGTCGCGAGACAGGTCGAGGAGGAGCATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGGTCCACGAGCGGAGGAGGAGTCTCTCGCCAGGCGCTAAC 363

RESULT 14
B0619318
LOCUS
DEFINITION
mays cDNA clone RNOSEQ6D01_SK.ab1 similar to No homology, mRNA
sequence.

ACCESSION B0619318
VERSION B0619318
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 648)
AUTHORS Wang, H. and Bohnert, H.J.
TITLE Genomics of plant stress tolerance

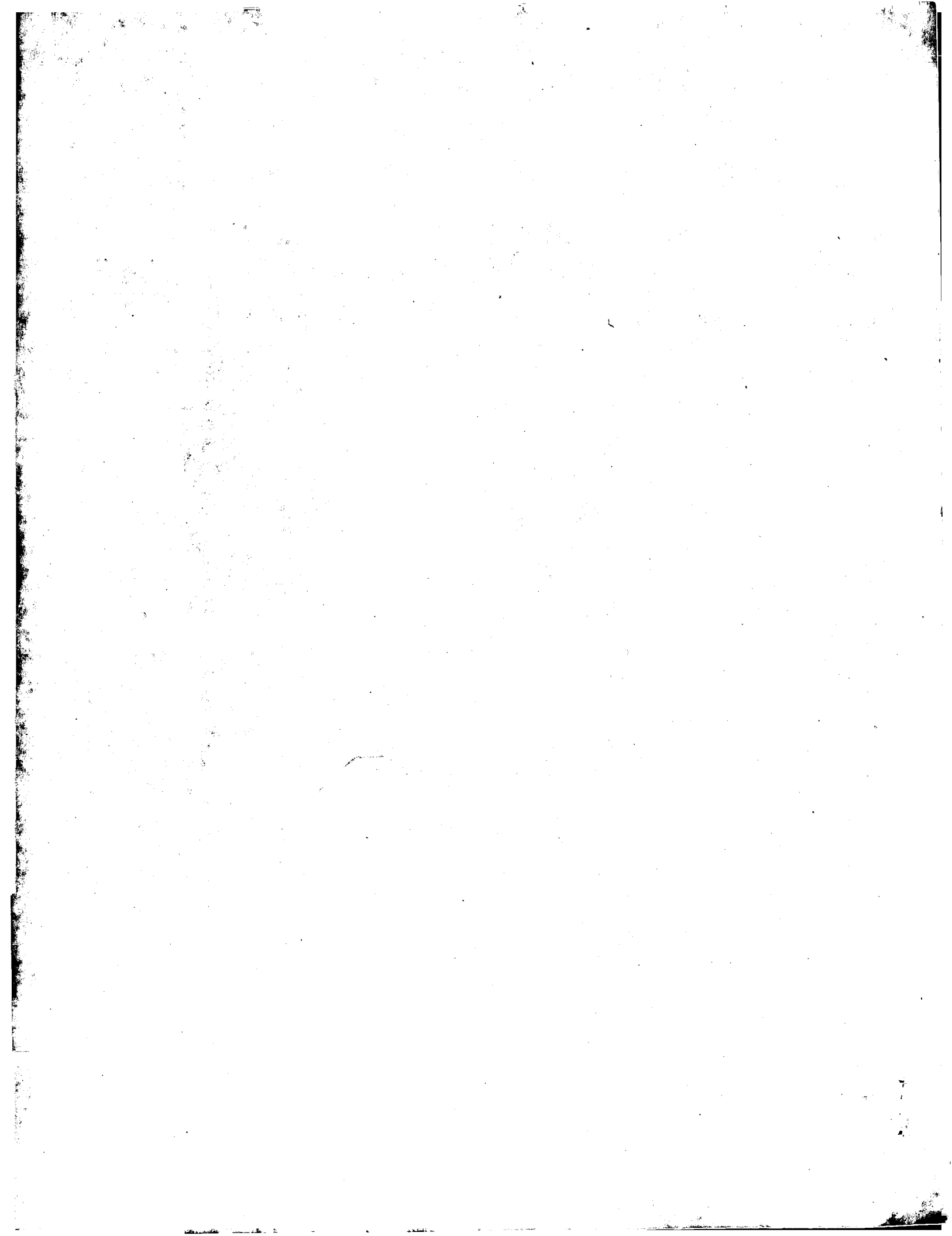
FEATURES Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
Source 1..648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RNOSEQ6E12_SK.ab1"
/tissue="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 1,33e-45 Length: 648
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: 14 Gaps: 1

US-10-090-035-4 (1-93) x B0619337 (1-648)

QY 1 MetalaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
Db 82 ATGGCTTACTACAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCCCGGCC 141
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
Db 142 GGCTTCGGCCGCCACGGAGCGGGTCCAGCAGCAGCGTCTCAAGGAGAGTTCGAGGAG 201
QY 41 ValAspThrValSerArgAlaGlyAlaAsn---HisHisHisHisHisGlyHisHisGly 59
Db 202 GTCGACACGGTCTCACGGCCCGCGCCACCCACCACCACCACCATGGTCAACACGGC 261
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGly 79
Db 262 GGCCACGGCTTCGTGTCGGCGAGACCAGGTCGAGGAGGAGACATCAACACCTGCACCGGC 321
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 322 GAGTCCACGAGCCGAGGAGAGCTTCCTCGCCAGGCTAAC 363

Search completed: June 16, 2003, 12:16:36
Job time : 1296.5 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 01:23:01 ; Search time 1222.39 Seconds
(without alignments)
7604.943 Million cell updates/sec

Title: US-10-090-035-3

Perfect score: 574

Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.2	90.8	553	11	AX104409
2	518	90.2	552	10	BE025302
3	509.6	88.8	546	9	AA979839
4	507.2	88.4	524	10	AW288876
5	500.8	87.2	523	10	AW331212
6	500.6	87.2	648	14	BQ619167

7	500.6	87.2	648	14	BQ619315
8	500.6	87.2	648	14	BQ619318
9	500.6	87.2	648	14	BQ619337
10	500.6	87.2	648	14	BQ619383
11	500.6	87.2	648	14	BQ619390
12	494.6	86.2	550	12	BG840383
13	494.4	86.1	539	9	AI855425
14	485.8	84.6	500	10	AW787732
15	465.6	81.1	495	10	AW787731
16	461.8	80.5	465	10	BE129897
17	457.2	79.7	476	10	BE129897
18	455.4	79.3	481	9	AI951929
19	440.4	76.7	466	9	AI964534
20	431.2	75.1	436	10	AW787315
21	426.2	74.3	443	10	BE129644
22	417	72.6	417	13	BM501439
23	411.4	71.7	433	10	BE225008
24	362.8	63.2	401	10	AW289056
25	326.6	56.9	541	10	AW745436
26	324.8	56.6	537	10	AW745400
27	322.8	56.2	326	10	BE025303
28	308.2	53.7	516	10	BE364814
29	308	53.7	519	10	AW680016
30	307.2	53.5	485	10	BE599123
31	307.2	53.5	513	14	BQ280709
32	304	53.0	522	10	AW679969
33	303.6	52.9	509	14	BQ280894
34	291.4	50.8	466	10	BE597738
35	274	47.7	362	12	BG840882
36	272.2	47.4	311	12	BF729420
37	251.4	43.8	274	10	AW288875
38	242	42.2	367	10	AW747603
39	240.6	41.9	399	13	BM318672
40	220.4	38.4	348	10	BE593507
41	219.6	38.3	226	9	AI372183
42	211	36.8	288	9	AI649804
43	208.8	36.4	338	10	AW924579
44	186	32.4	240	14	W21677
45	185.4	32.3	512	14	BQ283209

ALIGNMENTS

RESULT 1
AX104409
LOCUS AY104409 553 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PC0124784 mRNA sequence.
ACCESSION AY104409
VERSION AY104409.1 GI:21207487
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS

1 (bases 1 to 553)
Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

TITLE

Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes

JOURNAL

REFERENCE

2 (bases 1 to 553)
Coe, E.C.

TITLE

Direct Submission

JOURNAL

Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA

FEATURES

source

1.553
/organism="Zea mays"
/db_xref="MaizeDB:637271"
/db_xref="taxon:4577"
/clone="PC0124784"


```

FEATURES
Location/Qualifiers
1. 648
/organism="Zea mays"
/db_xref="taxon:4577"

```

```

FEATURES
    source
        Location/Qualifiers
            1. .523
                /organism="Zea mays"
                /cultivar="W23"
                /db_xref="taxon:4577"
                /clone_lib="707 - Mixed adult tissues from Walbot lab (SK
                )"
                /tissue_type="tassel, kernel, silk, husk, root, leaf"
                /dev_stage="adult"
                /lab_host="DH10B"
                /note="Organ: tassel, kernel, silk, husk, root, leaf;
                Vector: pGAD10; Site:1: ECORI; cDNA library from fully
                differentiated maize tissues from an active Mutator
                plant. Tissue ratio is 4/2/1/1/1 (tassel kernel, silk,

```


ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 648)
Wang,H. and Bohnert,H.J.
Genomics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172635473
Email: bohnertlab@life.uiuc.edu.

FEATURES
source
1..648
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone="RNOSEQ6D01_SK.abl"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /tissue_type="Roots"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150
 mM NaCl"

	BASE COUNT	172 a	193 c	183 g	100 t
ORIGIN					
Query Match		87.2%	Score 500.6;	DB 14;	Length 648;
Best Local Similarity		97.6%;	Pred. No. 1.8e-65;		
Matches 530;		Conservative 0;	Mismatches 9;	Indels 4;	Gaps 2;
QY	36	CACACAGCATTGCGAGTTCACACGGG-CGCTGCCGCACAGACACACCAAGCGTGCGCAC	94		
Db	21	CACGAGAGCATTTGAGCTGCGACGTCCGAGGGCGCTGCACACAGACACACCAAGCGTGCGCAC	80		
QY	95	AATGGCTTACTACACGAGGTGACACTACTGCTCGGAGGAGTGAGTGCTGGTGGCCCCGGC	154		
Db	81	AATGGCTTACTACACGAGGTGACACTACTGCTCGGAGGAGTGAGTGCTGGTGGCCCCGGC	140		
QY	155	CGGCTTGGCGGCGCACGGCGGGCGGTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGA	214		
Db	141	CGGCTTGGCGGCGCACGGCGGGCGGTCCAGCAGCAGCTGCTCAAGGAGAAGTTCGAGGA	200		
QY	215	GTCACACAGCTCTCACGGCGCGGGCGCA---ACCACCACCAACCATGGTTCACACGG	271		
Db	201	GGTCACACAGCTCTCACGGCGGGCGCAACACACACCAACCAACCATGGTTCACACGG	260		
QY	272	CGGCGCACGGCTTCGTGTGCGCAGACACAGGGTTCGAAGAGGACATCAACACCTGCACCG	331		
Db	261	CGGCGCACGGCTTCGTGTGCGCAGACACAGGGTTCGAAGAGGACATCAACACCTGCACCG	320		
QY	332	CGAGGTCCACAGCGCAGGAGGAGCTTCCTCCGACGGGTAACTGAGCGCGCGCGCGGC	391		
Db	321	CGAGGTCCACAGCGCAGGAGGAGCTTCCTCCGACGGGTAACTGAGCGCGCGCGCGGC	380		
QY	392	GGCATCCACGCGCGTTCGTGTTCGCTTCGCTTCGCTTATGTATGTCTGTGTTGACTGGTT	451		
Db	381	GGCATCCACGCGCGTTCGTGTTCGCTTCGCTTATGTATGTCTGTGTTGACTGGTT	440		
QY	452	GTTCAGGCTCATGCTACTTGGCTATCTAGTCGCACGCACTCAGCTCCCTGTAGCAATPAC	511		
Db	441	GTTCAGGCTCATGCTACTTGGCTATCTAGTCGCACGCACTCAGCTCCCTGTAGCAATPAC	500		
QY	512	GACAATAAGCTCGTGACCTGAATAAATCTTCTCGTAATACTAAAAAAAATAAAAAA	571		
Db	501	GACATAAGCTCGTGACCTGAATAAATCTTCTCGTAATACTAAAAAAAATAAAAAA	560		
QY	572	AAA 574			
Db	561	AAA 563			

Db	501	GACAATAGCTGCTGACCTGAATAAACTCTCTCGTAATACTAAAAA	563	
QY	572	AAA	574	
Db	561	AAA	563	
RESULT 10				
BQ619383				
LOCUS				
DEFINITION				
ACCESSION	BQ619383	648 bp	mRNA	linear
VERSION	BQ619383.1	GI:21621377		
KEYWORDS	EST			
SOURCE	zea mays			
ORGANISM	zea mays			
REFERENCE				
AUTHORS	Wang, H. and Bohnert, H.J.			
TITLE	Genomics of plant stress tolerance			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 217/2655473 Email: bohnerlab@life.uiuc.edu.			
FEATURES				
source				
	1..648			
	/organism="zea mays"			
	/db_xref="taxon:4577"			
	/clone="RNOSEQ7B02_SK.abl"			
	/clone_lib="Salt stressed zea mays roots cDNA library"			
	/tissue_type="Roots"			
	/dev_stage="2 weeks old"			
	/note="vector: pbluescript SK+; Stressed 24 hours at 150 mM NaCl"			
BASE COUNT	172 a	193 c	183 g	100 t
ORIGIN				
Query Match		87.2%	Score 500.6;	DB 14; Length 648;
Best Local Similarity		97.6%;	Pred. No. 1.8e-65;	
Matches	530;	Conservative	0;	Mismatches 9; Indels 4; Gaps 2;
QY	36	CACACAGCACTTCGAGGTCACAGGG-CGCTCGGCACACACACACACACAGCGTCGGCACC	94	
Db	21	CACGAGAGCACTTCGAGGTCGACGGCGCGCTCCACACACACACACACAGCGTCGGCACC	80	
QY	95	AATGGCTTACTACGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCCCGGC	154	
Db	81	AATGGCTTACTACGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCCCGGC	140	
QY	155	CGGCTTCGGCGCCACAGCGGGCGGCTCCAGCAGCAGCGTTCGAGGAGAGTTCGAGGA	214	
Db	141	CGGCTTCGGCGCCACAGCGGGCGGCTCCAGCAGCAGCGTTCGAGGAGAGTTCGAGGA	200	
QY	215	GGTCGACACGGTCTCAGCGCGCGCGCA-ACCACACACACACACACACACACACACAG	271	
Db	201	GGTCGACACGGTCTCAGCGCGCGCGCA-ACCACACACACACACACACACACACACAG	260	
QY	272	CGGCCACGGCTTCGTTGGTGGCGGAGACACAGGGTCGAGGAGAGATCAACACCTGCACCG	331	
Db	261	CGGCCACGGCTTCGTTGGTGGCGGAGACACAGGGTCGAGGAGAGATCAACACCTGCACCG	320	
QY	332	CGAGGTCACAGCGCGGAGGAGGTTTCCTCGCCAGGGCTAATGAGCGCCCGCGCGCC	391	
Db	321	CGAGGTCACAGCGCGGAGGAGGTTTCCTCGCCAGGGCTAATGAGCGCCCGCGCGCC	380	

resulting molecules were directionally cloned into the EcoRI and NotI sites of the p7/T3PAC vector. The library then went through one round of normalization to Cor value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

BASE COUNT 124 a 174 c 159 g 92 t 1 others

Query Match 86.2%; Score 494.6; DB 12; Length 550;
Best Local Similarity 97.6%; Pred. No. 1.5e-64;
Matches 524; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

QY	36	CACACAGCAGCTTCGACGTCACACGG-GCGCTGCGCACACACACCAAGCGTCGCGACC	94
DB	14	CCCAAGACCTTCGACGTCGCGGCGCTGCACACACACCAAGCGTCGCGACC	73
QY	95	AATGGCTTACTACACAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGCGCGGC	154
DB	74	AATGGCTTACTACACAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGCGCGGC	133
QY	155	CGGCTTCGCGCGCACGCGGCGGCTCCAGCAGCAGCTGCTCAAGAGAGTTTCGAGGA	214
DB	134	CGGCTTCGCGCGCACGCGGCGGCTCCAGCAGCAGCTGCTCAAGAGAGTTTCGAGGA	193
QY	215	GGTCGACACGGTCTCAGCGCGCGGCCA---ACCACCACACACCATGTCACACGG	271
DB	194	GGTCGACACGGTCTCAGCGCGCGGCCA---ACCACCACACACCATGTCACACGG	253
QY	272	CGCCACAGCTTCGCGGCGGAGACAGGCGTCAAGAGACATCAACACCTGCACCGG	331
DB	254	CGCCACAGCTTCGCGGCGGAGACAGGCGTCAAGAGACATCAACACCTGCACCGG	313
QY	332	CGAGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGCTAACTGAGCGCGCGCGGCC	391
DB	314	CGAGTCCACGAGCGCAGGAGAGCTTCTCGCCAGGCTAACTGAGCGCGCGCGGCC	373
QY	392	GGCATCCAGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	451
DB	374	GGCATCCAGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	433
QY	452	GTTACAGGTCATCGTACTTGGCTATCGTACGTCAGCGCAGCTCCTCTGTACGAAATAC	511
DB	434	GTTACAGGTCATCGTACTTGGCTATCGTACGTCAGCGCAGCTCCTCTGTACGAAATAC	493
QY	512	GACATAAGCTCGTGACCTGAATAAACTTCCTTCGTAATACATAAAAAA	568
DB	494	GACATAAGCTCGTGACCTGAATAAACTTCCTTCGTAATACATAAAAAA	550

RESULT 13
AI855425/c 539 bp mRNA linear EST 15-JUL-1999
LOCUS 603016F02.xl 603 - stressed root cDNA library from Wang/Bohnert lab
DEFINITION Zea mays cDNA, mRNA sequence.
ACCESSION AI855425
VERSION AI855425.1 GI:5499558
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 539)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Contact: Walbot V
Unpublished (1999)
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221

QY	272	CGGCCACGCTTCGTCGGCGGAGACAGGTCGACGACATCAACCTGCACCG	331
DB	261	CGGCCACGCTTCGTCGGCGGAGACAGGTCGACGACATCAACCTGCACCG	320
QY	332	CGAGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCGCGGCC	391
DB	321	CGAGTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCTAACTGAGCGCGCGGCC	380
QY	392	GGCATCCAGCGCGCTTCGTCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	451
DB	381	GGCATCCAGCGCGCTTCGTCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	440
QY	452	GTTACAGGTCATCGTACTTGGCTATCGTACGTCAGCGCAGCTCCTCTGTACGAAATAC	511
DB	441	GTTACAGGTCATCGTACTTGGCTATCGTACGTCAGCGCAGCTCCTCTGTACGAAATAC	500
QY	512	GACATAAGCTCGTGACCTGAATAAACTTCCTTCGTAATACATAAAAAA	571
DB	501	GACATAAGCTCGTGACCTGAATAAACTTCCTTCGTAATACATAAAAAA	560
QY	572	AAA 574	
DB	561	AAA 563	

RESULT 12
BG840383 550 bp mRNA linear EST 29-MAY-2001
LOCUS MEST12-H11.T7-1 ISUM4-TN Zea mays cDNA clone MEST12-H11 5', mRNA
DEFINITION sequence.
ACCESSION BG840383
VERSION BG840383.2 GI:14242676
KEYWORDS EST.
SOURCE Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 550)
Qiu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
On May 25, 2001 this sequence version replaced gi:14206705.
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu

PCR Primers
FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T7-1 (AA TAC GAC TCA CTA TAG).
Location/Qualifiers
1. .550
/organism="Zea mays"
/lab="host=DH10B"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST12-H11"
/clone_lib="ISUM4-TN"
/tissue_type="Seedling and silk"

ds-cDNA molecules were generated as follows. First-strand
cDNA was prepared from oligo-dT selected mRNA by priming
with a NotI oligo-dT primer (5'-
AAGTGAAGATTCGCGCGCAGGATTTTCTTTTCTTTT). The
resulting DNA:RNA hybrid was treated with RNase H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The

FEATURES source
Email: walbot@stanford.edu
Plate: 603016 row: F column: 02.
Location/Qualifiers
1. .539
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="603 - stressed root cDNA library from Wang/Bohnert lab"
/tissue_type="seedling"
/dev_stage="salt stress"
/lab_host="E. coli XL Gold"
/note="organ: root; Vector: pBluescriptII SK(+); XR; Seedling stressed root cDNA library from Wang/Bohnert lab"
BASE COUNT 91 a 157 c 172 g 119 t
ORIGIN

Query Match 86.1%; Score 494.4; DB 9; Length 539;
Best Local Similarity 97.2%; Pred. No. 1.6e-64;
Matches 525; Conservative 0; Mismatches 11; Indels 4; Gaps 2;
QY 34 TCCACACAGCACTTCGACGCTACACGGCGCTGCGCACACACACACACACGCTCGGCAC 93
Db TCCACACAGCACTTCGACGCTGCGACGGCGCTGCGCACACACACACACACGCTCGGCAC 480
QY 94 CAATGGCTTACTACACAGGAGTGGACTACTGCTCGGAGGAGGTGAGGTGCTGCGCCCGG 153
Db CAATGGCTTACTACACAGGAGTGGACTACTGCTCGGAGGAGGTGAGGTGCTGCGCCCGG 420
QY 154 CCGGCTTCGGCCGACCGCGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAAGTTCGAGG 213
Db CCGGCTTCGGCCGACCGGAGCGCGCTCCAGCAGCAGCTGCTCAAGGAGAAAGTTCGAGG 360
QY 214 AGGTCGACACGCTCTACGCGCGCGCGCGCA---ACCACACACACACACATCAACACCTGCACCG 270
Db AGGTCGACACGCTCTACGCGCGCGCGCGCA---ACCACACACACACACATCAACACCTGCACCG 300
QY 271 CGCGCCACGGGTTGCTGGTGGCGGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCG 330
Db CGCGCCACGGGTTGCTGGTGGCGGAGACCGAGGTCGAAGAGGACATCAACACCTGCACCG 240
QY 331 CGGAGGTCACGACCGGAGGAGGAGTTCCTCGCCAGGCTTAAGTACGCGCGCGCGCGG 390
Db CGGAGGTCACGACCGGAGGAGGAGTTCCTCGCCAGGCTTAAGTACGCGCGCGCGCGG 180
QY 391 CGGCATCCAGCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
Db CGGCATCCAGCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 451 TGTTCAGGTCATCGTACTTGGGTATCGTACGTGCAGCAGCTCAGCTCCTGTCGTAATTA 510
Db TGTTCAGGTCATCGTACTTGGGTATCGTACGTGCAGCAGCTCAGCTCCTGTCGTAATTA 61
QY 511 CGACAATAAGCTCTGACCTGAATAAAGCTCTCTGTAATTAATAAATAAATAAATAA 570
Db CGACAATAAGCTCTGACCTGAATAAAGCTCTCTGTAATTAATAAATAAATAAATAA 1

RESULT 14
AW787732
LOCUS AW787732 500 bp mRNA linear EST 16-MAY-2000
DEFINITION 945002E06.X3 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787732
VERSION AW787732.1 GI:7844510
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 500)
AUTHORS Walbot, V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. .500
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf; Vector: pGAD10; Site: 1; EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."
BASE COUNT 97 a 159 c 156 g 88 t
ORIGIN

Query Match 84.6%; Score 485.8; DB 10; Length 500;
Best Local Similarity 99.6%; Pred. No. 3.2e-63;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 62 GCCTCGGCACACACACACACGCTCGGCACCAATGGCTTACTACACAGGAGTGACCTA 121
Db GCCTCGGCACACACACACACGCTCGGCACCAATGGCTTACTACACAGGAGTGACCTA 60
QY 122 CTGCTCGGAGGAGGTGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
Db CTGCTCGGAGGAGGTGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 182 CCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAGGTTCGACACGCTCTACGCGCGCGCGC 241
Db CCAGCAGCAGCTGCTCAAGGAGAGTTCGAGGAGGTTCGACACGCTCTACGCGCGCGCGC 180
QY 242 CAACACACACACACACATGCTCACCACGCGCGCCACGCTTCTGCTGCTGCTGCTGCTGCTGCT 301
Db CAACACACACACACACATGCTCACCACGCGCGCCACGCTTCTGCTGCTGCTGCTGCTGCTGCT 240
QY 302 GGTGGAAGGAGCATCAACCTGCAACCGCGGAGGTCCACGAGCGCGGAGGAGCTTCCT 361
Db GGTGGAAGGAGCATCAACCTGCAACCGCGGAGGTTCACGAGCGCGGAGGAGCTTCCT 300
QY 362 CGCAGGCGCTAAGTACG 421
Db CGCAGGCGCTAAGTACG 360
QY 422 TGCCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
Db TGCCTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 482 GTGCAGGCACTCAGCTCCTGTACGAATTAAGCAATAAGCTCGTACCTGAATAAAGCTT 541
Db GTGCAGGCACTCAGCTCCTGTACGAATTAAGCAATAAGCTCGTACCTGAATAAAGCTT 480
QY 542 CTTTCGTAAT 550
Db CTTTCGTAAT 489
RESULT 15
AW787314

LOCUS AW787314 496 bp mRNA linear EST 16-MAY-2000
DEFINITION 945002E06.X1 945 - Mixed adult tissues from Walbot lab, same as 707
(SK) Zea mays cDNA, mRNA sequence.

ACCESSION AW787314
VERSION AW787314.1 GI:7844111
KEYWORDS EST.
SOURCE Zea mays.

ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 496)
AUTHORS Walbot.V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers

FEATURES
source
1..496
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

BASE COUNT 97 a 156 c 153 g 90 t
ORIGIN

Query Match 81.1%; Score 465.6; DB 10; Length 496;
Best Local Similarity 99.0%; Pred. No. 3.2e-60;
Matches 479; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 68 CGCACAGACACACACGCTCGGCACCAATGGCTTACTACGAGGAGTGACTACTGCTC 127
|||||
DB 1 CGCACAGACACACACGCTCGGCACCAATGGCTTACTACGAGGAGTGACTACTGCTC 60
|||||

QY 128 GGAGGAGTGAGTGCCTGCGCCCGCGGCTTTCGGCCGCCACGCGCGGCGTCCAGCA 187
|||||
DB 61 GGAGGAGTGAGTGCCTGCGCCCGCGGCTTTCGGCCGCCACGCGCGGCGTCCAGCA 120
|||||

QY 188 GCACGTGCTCAAGAGAGTTCAGAGAGTTCAGAGGTCGACAGGTCACGCGCGCGGCCAA-CC 246
|||||
DB 121 GCACGTGCTCAAGAGAGTTCAGAGAGTTCAGAGGTCGACAGGTCACGCGCGCGGCCAAATTC 180
|||||

QY 247 ACCACACACACACACGCTCGGCACGCGCGGCTTCGCTGTCGCGAGACACAGGTCG 306
|||||
DB 181 ATCACACACACACACGCTCGGCACGCGCGGCTTCGCTGTCGCGAGACACAGGTCG 240
|||||

QY 307 AAGAGGACATCAACACCTGCACCGGAGGTTCACAGGCGCAGGAGAGGTTCTCGCCA 366
|||||
DB 241 AAGAGGACATCAACACCTGCACCGGAGGTTCACAGGCGCAGGAGAGGTTCTCGCCA 300
|||||

QY 367 GGGCTAACTAGCGCGCGCGCGGCGCATCCACGCGGCTTCGCTGCTGCTGCGTGCCT 426
|||||
DB 301 GGGCTAACTAGCGCGCGCGCGGCGCATCCACGCGGCTTCGCTGCTGCTGCGTGCCT 360
|||||

QY 427 TATGTATGCTGTGTTGACTGTTGTTACAGGTCATCGTACTTGCGTATCGTACGTGCA 486
|||||

DB 361 TATGTATGCTGTGTTGACTGTTGTTGTCAGGGTTCATCGTACTTGGCTATCGTACGTGCA 420
QY 487 CGCACTCAGCTCCTGTGTACGAATTACGACAATAAGCTGCTGACCTGAATAAACTTCTTCG 546
|||||
DB 421 CGCACTCAGCTCCTGTGTACGAATTACGACAATAAGCTGCTGACCTGAATAAACTTCTTCG 480
QY 547 TAAT 550
|||||
DB 481 TAAT 484

Search completed: June 16, 2003, 03:55:52
Job time : 1228.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM protein - protein search, using sw model
Run on: June 6, 2003, 12:47:53 ; Search time 50 Seconds
(without alignments)
383.248 Million cell updates/sec
Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQVDYCEEVRSVAPA.....INTCTGEVHERRESFLARAN 93
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SPTREMBL_21:.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertibrate.*
14: sp.unclassified.*
15: sp.rviris.*
16: sp.bacteriap.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	274	53.7	229	10	Q8S0B1
2	177	34.7	95	10	Q9SM40
3	141	27.6	102	10	Q8S0B2
4	92.5	18.1	196	10	Q93ZA6
5	92.5	18.1	196	10	Q9SM63
6	92.5	18.1	259	10	Q2A643
7	89	17.5	1245	3	Q9P543
8	86.5	17.0	408	10	Q8S0B4
9	86.5	17.0	471	5	Q9UAY0
10	86	16.9	793	4	Q9Y2L9
11	84	16.5	441	10	Q944Q7
12	84	16.5	473	10	Q9FZ33
13	84	16.5	1561	5	Q9W3D2
14	83	16.3	168	5	Q9VWM5
15	82.5	16.2	381	13	Q9OWV0
16	82.5	16.2	735	5	Q9NES7

17	82	16.1	173	10	Q9LFS9
18	81	15.9	86	5	Q20691
19	81	15.9	109	11	Q9D6B9
20	81	15.9	191	10	Q42448
21	81	15.9	397	5	Q9NC84
22	80.5	15.8	347	5	Q9VX61
23	80.5	15.8	396	10	Q9FH24
24	80	15.7	133	10	Q9ZR12
25	80	15.7	1273	4	O60316
26	79	15.5	612	10	Q941S9
27	79	15.5	612	10	Q8RUM5
28	78	15.3	156	10	Q39754
29	78	15.3	585	5	Q9U4F0
30	78	15.3	585	5	Q8T8F8
31	78	15.3	990	5	Q9U4F1
32	78	15.3	1020	5	Q9U1K1
33	77.5	15.2	489	17	O8THX7
34	77.5	15.2	537	5	Q9VTJ6
35	77.5	15.2	566	5	Q9VTJ7
36	77.5	15.2	709	3	Q9VW52
37	77.5	15.2	1693	5	Q9VXV8
38	77	15.1	354	5	O8SSW8
39	77	15.1	509	5	Q94888
40	77	15.1	509	5	Q9VZN8
41	77	15.1	522	5	O9XTK9
42	77	15.1	834	6	O9GKQ7
43	76.5	15.0	77	5	Q20690
44	76.5	15.0	332	16	Q9KG54
45	76.5	15.0	1283	5	Q95RH4

ALIGNMENTS

RESULT 1
Q8S0B1 PRELIMINARY; PRT; 229 AA.
ID Q8S0B1
AC Q8S0B1;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DE P0470A12.5 protein.
DE P0470A12.5
GN P0470A12.5
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12.5";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003436; BAB90280.1;
SQ SEQUENCE 229 AA; 25296 MW; D6ED4AA5FFEF61E CRC64;

Query Match 53.7%; Score 274; DB 10; Length 229;
Best Local Similarity 58.1%; Pred. No. 4.9e-23;
Matches 54; Conservative 14; Mismatches 21; Indels 4; Gaps 2;
QY 2 AYQVDYCEEVRSVAP-AGFGRHGGVQOHHVKEFEEDVTYSRAGANHHHHHHGG 60
Db 140 AQEKEVDYCEEVRSVTPGFLGRGGVQOHHVKEFEEDVTYSRAGANHHHHHHGG 196
QY 61 HGFVRETVEEDINTCTGEVHERRESFLARAN 93
Db 197 DYLMVRETVEEDINTCTGEVHERRESFLARAN 229
RESULT 2
Q9SM40

Mon Jun 16 14:55:59 2003

```

01-JUN-2002 (TREMBlrel. 21, Last annotation update)
AT4g36900/C7A10.460.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
Bowler L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.; Arabidopsis cDNA clones.
Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
EMBL: AY057683; AAL1514.1;
InterPro: IPR001471; TF_ERF.
Pfam: PF00847; AP2-domain; 1.
ProDom: PD001423; TF-AP2; 1.
SEQUENCE 196 AA; 21392 MW; E423D5570745C9CD CRC64;
Query Match 18.1%; Score 92.5; DB 10; Length 196;
Best Local Similarity 33.8%; Pred. No. 0.013;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;
QY 21 GFGRHGGG-VQOHVKEKFE-----VDTVSRAGA---NHHHHGH-HGGHGVVRET--R 69
| | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db 99 GGGVNGGDSAAVIRKAAEVAQVDALEAAGAGNRRHHHHHHHGHGNDYDNRSDYR 158
| | | | | : | | | | | | | | | | | | | | | | | | | | | |
QY 70 VEEDINTCTGE 80
| | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db 159 INDDLMECSK 169
| | | | | : | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
Q9SM63 PRELIMINARY; PRT; 196 AA.
AC Q9SM63
DT 01-JUN-2002 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE TIN-like protein.
GN C7A10.460 OR AT4G36900.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Mayer K.F.X.,
Schueller C.
Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
EU Arabidopsis sequencing project;
Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
EU Arabidopsis sequencing project;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z99707; CAB16766.1;
DR EMBL: AL161590; CAB80356.1;
DR HSSP: O80337; 2GCC.
DR InterPro: IPR001471; TF_ERF.
DR Pfam: PF00847; AP2-domain; 1.
DR PRINTS: PR00367; ETHRSPLEMT.
DR ProDom: PD001423; TF-AP2; 1.
DR SMART: SM00380; AP2; 1.

Q9SM40 PRELIMINARY; PRT; 95 AA.
AC Q9SM40;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-WAR-2001 (TREMBlrel. 16, Last annotation update)
DE Putative glycine-rich protein.
OS Sporobolus stapfianus (Resurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Eragrostideae; Sporobolus.
NCBI_TaxID=56623;
[1]
SEQUENCE FROM N.A.
Neale A.D., Blomstedt C.K., Bronson P., Le T.N., Guthridge K.,
Evans J., Gaff D.F., Hamill J.D.;
"The isolation of lowly-transcribed genes which are induced during
RT deacclimation of the resurrection grass Sporobolus stapfianus."
Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AJ242802; CAB61838.1;
SEQUENCE 95 AA; 10668 MW; D756DCE2B68DD85B CRC64;
Query Match 34.7%; Score 177; DB 10; Length 95;
Best Local Similarity 57.7%; Pred. No. 1.5e-12;
Matches 45; Conservative 8; Mismatches 15; Indels 10; Gaps 6;
QY 1 MAYTQE-VDYCSSEVRSVAPAGFGRH-GGGVQOHVKEKFEVDVTSRAGANHHHHHHGH 58
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MAHFKEMVDY--EVTSMKPGRGHGGGVQGVVVKETFEVEVQVTPRGRS-GHHGHR 57
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 G--GHG---FVRETRVE 71
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 GNOGHGSGHFQARETFE 75
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
Q9S0B2 PRELIMINARY; PRT; 102 AA.
AC Q9S0B2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE P0470A12.4 protein.
GN P0470A12.4.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=39947;
[1]
SEQUENCE FROM N.A.
STRAIN-CV. NIPPONBARE;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0470A12."
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003436; BAB90279.1;
DR EMBL: 11214 MW; 6FF1266B1CDE7768 CRC64;
SEQUENCE 102 AA; 11214 MW;
Query Match 27.6%; Score 141; DB 10; Length 102;
Best Local Similarity 45.8%; Pred. No. 2e-08;
Matches 27; Conservative 8; Mismatches 18; Indels 6; Gaps 1;
QY 3 YYQEVDCSEVRSVAPAGFGRHGGGVQOHVKEKFEVDVTSRAGANHHHHHHHHGH 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 YYSEVDHCAEMNRPP-----HAGGEHVAVRSEYEEVDENARAGRHHHGGGGGH 56
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
Q93ZA6 PRELIMINARY; PRT; 196 AA.
AC Q93ZA6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

```

SQ      SEQUENCE      196 AA;  21362 MW;  0023D5571345C6A0 CRC64;

Query Match      18 1%;  Score 92.5;  DB 10;  Length 196;
Best Local Similarity 33.8%;  Pred. No. 0.013;
Matches 24;  Conservative 12;  Mismatches 24;  Indels 11;  Gaps 5

QY      21  GFGRHGGG-VQOHHVVVEKFEF---VDTVSAGA---NHHHHHHH-HGSHGFVRET--R 69
      | | | | | : | | | | | | | | | | | | | | | | | | | | |
Db      99  GGGVNGGCDMSAAVIRKAAEVGAQVDLEAAGAGGNRRHHHHHQQHQRGNHDYVDNHSYR 158

QY      70  VEEDINTCTGE 80
      : | : | :
Db      159  INDDLMCESSK 169

```

RESULT 6	PRELIMINARY;	PRT;	259 AA.
ID O24643			
AC O24643;			
DT 01-JAN-1998	(TREMBlrel. 05, Created)		
DT 01-JAN-1998	(TREMBlrel. 05, Last sequence update)		
DT 01-JUN-2002	(TREMBlrel. 21, Last annotation update)		
DE TINY-like protein (Fragment).			
OS Arabidopsis thaliana (Mouse-ear cress).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
NCBI_TaxId=3702;			
[1]	SEQUENCE FROM N.A.		
NRN STRAIN=COLUMBIA;			
RC Terryn N.;			
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.			
[2]			

MEDLINE-97338147; PubMed-9192694;
Okumuro J.K., Caster B., Villarroel R., van Montagu M., Jofuku K.D.;
"The AP2 domain of APTAL2 defines a large new family of DNA binding
proteins in Arabidopsis.",
Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).
ENBL; AFJ002598; CAA05630.1; -;
EMBL; AF003103; AAC49776.1; -;
HSP; O80337; 3GCC.
InterPro: IPR001471; TF_ERF.
Pfam; PF00847; AP2-domain; 1.
PRINTS; PR00367; ETHRSPELEMT.
PRODOM; PD001423; TF_AP2; 1.
SMART; SM00380; AP2; 1.
NON_TER 1 1
CHAIN 64 >259 TINY-LIKE PROTEIN.
NON_TER 259 259
SEQUENCE 259 AA; 28615 MW; 0464949DBSGC619DB CRC64;

Query Match	18.1%;	Score 92.5;	DB 10;	Length 259;
Best Local Similarity	33.8%;	Pred. No. 0.018;	24;	
Matches	24;	Conservative 12;	Mismatches 11;	Gaps 5;
y	21	GFGRHGGG-VQOHVVKERFEE----	VDTVSRAGA----	NNHHHHGH-HGGHCFVVRET--R 69
b	162	: : : : :	: : : : :	: : : : :
		GGVNGGGDMSAAYTRRAAEVGAQVDLEAAGAGGNNRHHHHHHHHOGRNHDYDNDHSDYR 221		
y	70	VEEDINTCTGE 80		
b	222	: : : : : : :		
		INDDLMEGSSK 232		

[illegible]

DE Related to multicyclin-dependent kinase PHO85.
GN B24H17.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxId=5141;
[1]
RN SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL356815; CAB92623.1; -.
DR HSP: P42773; IIIB...
DR InterPro: IPR002110; ANK.
DR InterPro: IPR004129; GDPD.
DR InterPro: IPR004331; SPX.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF03009; GDPD; 1.
DR Pfam: PF03105; SPX; 1.
DR SMART: SM00248; ANK; 3.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT; 1.
KW ANK repeat; Cyclin; Kinase; Repeat.
SQ SEQUENCE 1245 AA; 137798 MW; 16BD0F6A04596A9F CRC64;

Query Match	17.5%	Score 89;	DB 3;	Length 1245;
Best Local Similarity	72.2%;	Pred. No. 0.29;		
Matches 13;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;
QY	45	SRAGANHHHHHGGHGG 62		
		:		
Db	901	SRSGTHHHHHHHHGGHG 918		

```

RESULT 8
Q8S0B4
ID Q8S0B4 PRELIMINARY; PRT; 408 AA.
DT AC Q8S0B4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE P0470A12.2 protein.
GN P0470A12.2
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
NCBI_TaxID=39947;
[1]
RN
RX
OR
RC SEQUENCE FROM N. A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0470A12."
RT
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003436; BAB90277.1; -.
RS
SQ SEQUENCE 408 AA; 43466 MW; F760237877F6D69C CRC64;

```

```

Query Match      17.0%; Score 86.5; DB 10; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.15;
Matches 25; Conservative 10; Mismatches 15; Indels 43; Gaps 5;

12 EEVRSVAPAGFGRH-----GGVQVHHVKEKEPEEDVTSRAGA-----NHH----- 52
|| : | :| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
199 EECDAAGGGYGRHHGAGGAVKQHATYKQHQALTEGVNGCGAGYNNRQHAVAVAGGQ 258
|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

53 HHH-----GHH--GGH 61
|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

259 HHYGATAAAYGNASNKQHFTAAAGHHSSGH 291
|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

RESULT 9	
Q9UAY0	PRELIMINARY; PRT; 471 AA.
AC Q9UAY0;	
DT 01-MAY-2000 (TRENBLrel. 13, Created)	
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)	
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)	
DE W03G1.5 protein.	
GN W03G1.5.	
OS Caenorhabditis elegans.	
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;	
OC Rhabditidae; Peloderinae; Caenorhabditis.	
OX NCBI_TaxID=6239;	
[1]	
SEQUENCE FROM N.A.	
RP STRAIN-BRISTOL N2;	
RC MEDLINE=94150718; PubMed=7906398;	
RX Wilson R., Ainscough R., Connell M., Copsey T., Cooper J., Coulson A., Bonfield J., Burton J., Dunn J., Favello A., Fulton L., Craxton M., Dear S., Du Z., Durbin R., Hillier L., Jier M., Johnston L., Gardner A., Green P., Hawkins T., Hillier N., Latreille P., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R., Smalton N., Smith A., Sonnhammer E., Staden K., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.; "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans"; Nature 368:32-38(1994).	
RL Nature 368:32-38(1994).	
[2]	
SEQUENCE FROM N.A.	
RP STRAIN-BRISTOL N2;	
RC STRAIN-BRISTOL N2;	
RX Pauley A., Scheet P., Harper M.;	
RA "The sequence of C. elegans cosmid W03G1.5";	
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
[3]	
SEQUENCE FROM N.A.	
RP STRAIN-BRISTOL N2;	
RC Waterston R.;	
RA EMBL: AF125964: AAD14753.1;	
RL EMBL: AF125964: AAD14753.1;	
SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A6A4985B CRC64;	
Query Match 17.0%; Score 86.5; DB 5; Length 471;	
Best Local Similarity 34.8%; Pred. No. 0.17; 16; Indels 25; Gaps 4;	
Matches 24; Conservative 4; Mismatches 16; Indels 25; Gaps 4;	
QY 15 RSVAPAGFGRHGG-----GVQOHVVKKEFEVDVTSRAGANHHHHHGG-----56	
Db 372 RSHSPRGHG-HGGRHGPFGHGHGPPPH-----HHHGRSPSRHHHHHHHGGCRPF 426	
QY 57 --HRHGHHGF 63	
Db 427 PPHGHGHGF 435	
RESULT 10	
Q9V2L9	PRELIMINARY; PRT; 793 AA.
AC Q9V2L9;	
DT 01-NOV-1999 (TRENBLrel. 12, Created)	
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)	
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)	
DE KIAA1016 protein (Fragment).	
GN KIAA1016.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
[1]	

SEQUENCE FROM N.A.	
RP TISSUE-BRAIN;	
RC MEDLINE=99446063; PubMed=10231032;	
RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;	
RA "Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";	
RL DNA Res. 6:63-70(1999).	
DR EMBL: AB023233; BAA76860.1; in-like.	
DR InterPro: IPR001715; Calponin-like.	
DR InterPro: IPR002114; Hpr_Serp_site.	
DR InterPro: IPR001611; LRR.	
DR InterPro: IPR003592; LRR_out.	
DR InterPro: IPR003591; LRR_typ.	
DR Pfam: PF00307; CH; 1	
DR Pfam: PF00560; LRR; 6.	
DR PRINTS: PR00019; LEURICHRPT.	
DR SMART: SM00033; CH; 1.	
DR SMART: SM00370; LRR; 3.	
DR SMART: SM00369; LRR_TYP; 1.	
DR PROSITE: PS00021; CH; 1.	
DR PROSITE: PS00589; PTS_HPR_SER; UNKNOWN_1.	
FT NON_TER 1	
SQ SEQUENCE 793 AA; 87801 MW; 0BE99CC48C3B37C CRC64;	
Query Match 16.9%; Score 86; DB 4; Length 793;	
Best Local Similarity 32.8%; Pred. No. 0.37; 28; Indels 10; Gaps 2;	
Matches 22; Conservative 7; Mismatches 28; Indels 10; Gaps 2;	
QY 6 EVDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFE-----VDVTSRAGA-----NHHHHH 55	
Db 39 ELFPSPRGRCDDPPAGAGRGGRKXKATPGPQFPFVLSVATLPLHLHHHHHHH 98	
QY 56 GHGGHG 62	
Db 99 QHGGTG 105	
RESULT 11	
Q944Q7	PRELIMINARY; PRT; 441 AA.
AC Q944Q7;	
DT 01-DEC-2001 (TRENBLrel. 19, Created)	
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)	
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)	
DE Atlg54990/F14C21_5.	
OS Arabidopsis thaliana (Mouse-ear cress).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
OX NCBI_TaxID=3702;	
[1]	
SEQUENCE FROM N.A.	
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;	
RA "Arabidopsis cDNA clones.";	
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL: AF424608; AAL11602.1;	
SQ SEQUENCE 441 AA; 48816 MW; 520163FE0A8DE447 CRC64;	
Query Match 16.5%; Score 84; DB 10; Length 441;	
Best Local Similarity 36.8%; Pred. No. 0.31; 18; Indels 8; Gaps 3;	
Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;	
QY 8 DYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDVTSRAGANHHHHHGGH---HGH 61	
[1]	

Db 377 DYISEFV-SLLPKSIRR-----VAEPIPEEVQKVLEAKAGDDHDDHHHGCHAHAGY 428

RESULT 12

```
Q9FZ33
ID Q9FZ33 PRELIMINARY; PRT; 473 AA.
AC Q9FZ33;
DT 01-NAR-2001 (TrEMBLrel. 16, Created)
DT 01-NAR-2001 (TrEMBLrel. 16, Last sequence update)
DE T24C10.10 protein (Hypothetical 52.4 kDa protein).
DE T24C10.10 OR F14C21.51 OR AT1G54990.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC OC Spermatophyta; Magnoliophyta; eudicotyledons;
OC OC eudicotyledons; core eudicotyledons; Rosidae;
OC OC eudicotyledons; Brassicales; Brassicaceae; Arabidopsi.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altfati H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez S.,
RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV_COLUMBIA;
RA MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altfati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Eguo P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.S., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Mafti R., Marzilli A.,
RA Millitecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RA thaliana".
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamlaya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davila R.W., Ecker J.R., Theologis A.;
RA "Full length cDNA of gene At1G54990 (GI:15221965)".
RN Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RA EMBL; AC064840; AAC00878.1; -
RA EMBL; AC069144; AAC51115.1; -
RA EMBL; AY074334; AAL67030.1; -
RA Hypothetical protein.
RN SEQUENCE 473 AA; 52419 MW; D09124A11565BB23 CRC64;
```

```

Query Match      16.5%  Score 84;  DB 10;  Length 473;
Best Local Similarity 36.8%  Pred. No. 0.34;
Matches 21;  Conservative 10;  Mismatches 18;  Indels 8;  Gaps 3;

QY      8  DYCSEEVRSVAPAGCFGRHGGGVQOHHVVKFEEDVTVSAGANHHHHHH--HGHH 61
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

Db	409	DYISEFV-SLLPKSIRR-----VAEERPIEEVQKVL	EAKAGDDHHHHGHCHAHAGY	460
	RESULT 13			
	Q9W3D2			
ID	Q9W3D2	PRELIMINARY;	PRT; 1561 AA.	
AC	Q9W3D2;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	CRAG protein.			
GN	CRAG OR CGI2737.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			

[1] SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatzides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S.A., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gongtellan A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Globis R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
TX Science 287:2185-2195(2000).

```

RR EMBL: AE003444; AAF64397.1; -
RR FlyBase: FBgn0025864; Crag
RR InterPro: IPRO01112; dENN.
RR InterPro: IPRO01194; dENN.
RR InterPro: IPRO05113; uENN.
RR Pfam: PF03455; dENN; I.
RR Pfam: PF02141; dENN; I.
RR Pfam: PF03456; uENN; I.
RR Q SEQUENCE 1561 AA; 174616 MW; 10BF038A3DD4FC4E CRC64;

```

Query Match 16.5%; Score 84; DB 5; Length 1561;
Best Local Similarity 23.9%; Pred. No. 1.4;
Matches 28; Conservative 17; Mismatches 30; Indels 42; Gaps 5;

121
SEQUENCE FROM N.A.
RN
RP
RC STRAIN-BREKELY;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA George M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA Champe M., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Paciej J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA

	Query Match	16.2%	Score 82.5;	DB 13;	Length 381;
	Best Local Similarity Matches	29.0%;	Pred. No. 0.38;	Mismatches 9;	Indels 3; Gaps 1
	Conservative	18;			
QY	3 YVEVDYCSEVRS---	VAPAGFGRHGGVGQHVVKKEFEVDTVSRAGANHHHHHHGHC	59		
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Dd	57 FYESELCSANATPLTAPLKAGARHLNRQHQQOQQOHHQNNHHGHQ	116			
	:	:	:	:	:
	:	:	:	:	:
QY	60 GH 61				
Dd	117 DH 118				

Search completed: June 6, 2003, 12:52:28
Job time : 51 secs


```

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:      June 6, 2003, 12:48:33 ; Search time 25 Seconds
              (without alignments)
              357.620 Million cell updates/sec

Title:      US-10-090-035-4
Perfect score: 510
Sequence:    1 MAYTQEDVDCSEEVSVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table:  BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:      283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :      PIR_73:*
1:      pir1:*
2:      pir2:*
3:      pir3:*
4:      pir4:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	92.5	18.1	196	2	G85435	TINY-like protein
2	92.5	18.1	239	2	T52619	TINY-like protein
3	92	18.0	414	2	A48273	delta/YVI/NF-E3/UC
4	89	17.5	1245	2	T49815	related to multifu
5	86.5	17.0	471	2	T33997	hypothetical prote
6	84	16.5	473	2	D96591	hypothetical prote
7	82.5	16.2	735	2	T45059	hypothetical prote
8	82	16.1	173	2	T51469	glycine/proline-ri
9	81.5	16.0	495	1	S31223	transcription fact
10	81	15.9	86	2	T16437	hypothetical prote
11	81	15.9	191	2	A42844	abscisic acid-and
12	80	15.7	133	2	T09608	environmental stre
13	80	15.7	1273	2	T00338	hypothetical prote
14	79	15.5	204	2	T09592	protein corA, cold
15	79	15.5	414	2	A40350	transcription repr
16	77	15.1	754	2	JC4898	Down-syndrome-crit
17	76.5	15.0	77	2	T16436	hypothetical prote
18	76.5	15.0	332	2	C83682	hypothetical prote
19	76	14.9	102	2	T30119	hypothetical prote
20	76	14.9	1560	2	T00080	hypothetical prote
21	75.5	14.8	425	1	JH0710	transcription fact
22	75	14.7	369	1	TVFVAF	transforming prote
23	75	14.7	600	2	F71434	probable limonene
24	75	14.7	1019	2	T00117	dve protein - fru
25	75	14.7	1441	2	T13717	CRAG protein - fru
26	74.5	14.6	410	2	T26757	hypothetical prote
27	74	14.5	83	2	T16435	hypothetical prote
28	74	14.5	895	2	JC7089	zinc finger bindin
29	74	14.5	1891	2	T13594	hypothetical prote

Query Match 18.1%; Score 92.5; DB 2; Length 259;

ALIGNMENTS

RESULT 1
G85435

TINY-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G85435
 C:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold sp
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A95001; M0ID:20083488; PMID:10617198
 A:Accession: G85435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-196 <STO>
 A:Cross-references: GB:NC_001268; NID:g7270639; PIDN:CAB80356.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G36900
 A:Map position: 4

Query M
Best R

Matches

Qy

qQ

2y

q

RESULT 2

52619
YNY-like

Species: 24

Accession
Terry,

eller, C.
EBS Lett.

;Title: E
;Referenc

```

;Accesslo
;Status:

```

;Molecules
;Residues;Cross-re
;Genetics
;Man...

Quercus. Map post

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1245 <SCH>
A;Cross-references: EMBL:AL356815; GSPDB:GN00116; NCSP:B24H17.10
A;Experimental source: BAC clone B24H17; strain OR74A
C;Genetics:
A;Gene: NCSP:B24H17.10
A;Map position: 6
A;Introns: 6/2; 1141/3

Query Match 17.5%; Score 89; DB 2; Length 1245;
Best Local Similarity 72.2%; Pred. No. 0.21;
Matches 13; Conservative 2; Mismatches 0; Gaps 0;

QY 45 SRAGANHHHHGHHGGHG 62
Db 901 SRSGTHHHHHHHHHGGHG 918

RESULT 5
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33997
R;Pauley, A.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid W03G1.
A;Reference number: Z21454
A;Accession: T33997
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-471 <PAU>
A;Cross-references: EMBL:AF125964; PIDN:AA14753.1; GSPDB:GN00022; CESP:W03G1.5
A;Experimental source: strain Bristol N2; clone W03G1
C;Genetics:
A;Gene: CESP:W03G1.5
A;Map position: 4

Query Match 17.0%; Score 86.5; DB 2; Length 471;
Best Local Similarity 34.8%; Pred. No. 0.14;
Matches 24; Conservative 4; Mismatches 16; Indels 25; Gaps 4;

QY 15 RSVAPAGRGHGG-----GVQHVKEFEVDVTSRAGANHHHHG----- 56
Db 372 RSHSPRGHG-HGGRHPPHPCGRHGHGPHH-----HHHGKSPSRHHHHHHHGGCRPF 426

QY 57 --HHGGHG 63
Db 427 PPHGHGHF 435

RESULT 6
D96591
hypothetical protein T24C10.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96591
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>

Best Local Similarity 33.8%; Pred. No. 0.016;
Matches 24; Conservative 12; Mismatches 24; Indels 11; Gaps 5;

QY 21 GFRHGGG-VQOHVVKKEE-----VDTVSRAGA---NHHHHGH-HGGHGFVVRVET--R 69
Db 162 GGVNGGDMASAYIRRAAEVGAQVDALEAAGAGNRHHHHHHQHQRGNDHYVDNHSYR 221

QY 70 VEEDINTCTGE 80
Db 222 INDLMCCSSK 232

RESULT 3
A48273
delta/Y1/NF-EL/UCRBP transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C;Accession: A48273; A42055; A56418
R;Safrany, G.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 90, 5559-5563, 1993
A;Title: Characterization of the mouse gene that encodes the delta/Y1/NF-EL/UCRBP trans
A;Reference number: A48273; MUID:93296177; PMID:8516301
A;Accession: A48273
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-414 <RES>
A;Cross-references: GB:L13968; NID:g293847; PIDN:AAA0477.1; PID:g293849
R;Flanagan, J.R.; Becker, K.G.; Ennist, D.L.; Gleason, S.L.; Driggers, P.H.; Levi, B.Z.;
Mol. Cell. Biol. 12, 38-44, 1992
A;Title: Cloning of a negative transcription factor that binds to the upstream conserved
A;Reference number: A42055; MUID:92107191; PMID:1309593
A;Accession: A42055
A;Molecule type: mRNA
A;Residues: 1-414 <FLA>
A;Cross-references: GB:M73963; NID:g202270; PIDN:AAA0522.1; PID:g202271
A;Note: sequence extracted from NCBI backbone (NCBIN:74641, NCBI:74642)
R;Hartharan, N.; Kelley, D.E.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 9799-9803, 1991
A;Title: delta, a transcription factor that binds to downstream elements in several poly
A;Reference number: A56418; MUID:92052178; PMID:1946404
A;Accession: A56418
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218, 'S', 220-374, 'G', 376-414 <HAR>
A;Cross-references: GB:M74590; NID:g192940; PIDN:AAA37521.1; PID:g192941
A;Introns: 227/1; 281/2; 301/3; 354/3
C;Keywords: transcription factor; zinc finger

Query Match 18.0%; Score 92; DB 2; Length 414;
Best Local Similarity 31.0%; Pred. No. 0.031;
Matches 22; Conservative 7; Mismatches 26; Indels 16; Gaps 2;

QY 21 GFRHGGG-VQOHVVKKEEVDVTSRAGANHHHHHHHHGGHGGHGFVVRVETRVVEEDINTCTGE 80
Db 55 GGGDHGGGGGGH-----GHAGHHHHHHHHHHHHHPMTALQPLVTD-----PTQ 98

QY 81 VHERRESFLAR 91
Db 99 VHHHQEVILVQ 109

RESULT 4
T49815
related to multifunctional cyclin-dependent kinase PHO85 [imported] - Neurospora crassa
A;Alternate names: protein B24H17.10
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49815
R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: 225022
A;Accession: T49815

A:Cross-references: GB:AE0051173; NID:g9857523; PIDN:AG00878.1; GSPDB:GN00141
C:Genetics:
A:Gene: T24C10.10
A:Map position: 1

Query Match 16.5%; Score 84; DB 2; Length 473;
Best Local Similarity 36.8%; Pred. No. 0.26;
Matches 21; Conservative 10; Mismatches 18; Indels 8; Gaps 3;

QY 8 DYCSEEVRSVAPAGFRHGGVQVHVKEFEVDVTSRAGANHHHHHGHG---HGHH 61
DB 409 DYISEFV-SLLPKSIRR-----VAEPIPEVQKVLAEAKAGDDHDDHHGHGHAHAGY 460

RESULT 7

hypothetical protein Y39B6B.gg [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45059
R:Wilson, R.; Alnsough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.; Nature 368, 32-38, 1994
A:Authors: Showkeen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stock, L.; Wilkinson-Sproat, J.; Wohldman, P.
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID:94150718; PMID:7906398
A:Accession: T45059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-735 <WIL>
A:Cross-references: EMBL:AL132896; NID:g6434440; PIDN:CAB60938.1; PID:g6434473
A:Experimental source: clone Y39B6B
C:Genetics:
A:Map position: 3
A:Introns: 18/1; 69/1
A:Note: Y39B6B.gg

Query Match 16.2%; Score 82.5; DB 2; Length 735;
Best Local Similarity 26.4%; Pred. No. 0.6;
Matches 19; Conservative 1; Mismatches 9; Indels 43; Gaps 2;

QY 23 GRHGGVQVHVKEFEVDVTSRAGANHHH-----HGHHGAHHHAPHHEHGHGHHGHHGHTHH 662
DB 614 GHGAGYGAH-----HGHHGAHHHAPHHEHGHGHHGHHGHTHH 662
QY 54 ---HHGHHGGHG 62
DB 663 SLAHGHHGGHG 674

RESULT 8

glycine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein K10A8_130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51469
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <SAT>
A:Cross-references: EMBL:AL391151
A:Experimental source: cultivar Columbia; BAC clone K10A8
C:Genetics:
A:Map position: 5
A:Introns: 97/1
A:Note: K10A8_130

Query Match 16.1%; Score 82; DB 2; Length 173;
Best Local Similarity 34.7%; Pred. No. 0.15;
Matches 17; Conservative 7; Mismatches 15; Indels 10; Gaps 2;

QY 19 PAGFRHGGVQVHVKEFEVDVTSRAGA---NHHHHHHHHGGHGF 63
DB 93 PSHSGHHGGTGATGAG-----GVAAGAAGAHMHHHGHGHHGHG 135

RESULT 9

S31223
transcription factor Brn-1 - mouse
N:Alternate names: class III POU domain protein brain-1
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
C:Accession: S31223
R:Hara, Y.; Roverscalli, A.C.; Kim, Y.; Mirenborg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A:Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A:Reference number: S31223; MUID:92228768; PMID:1565620
A:Accession: S31223
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAR>
A:Cross-references: EMBL:M88299; NID:g200444; PIDN:AAA39960.1; PID:g200445
C:Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:27-49/Region: glycin-rich
F:101-112/Region: alanine-rich
F:162-180/Region: histidine/proline-rich
F:186-201/Region: alanine-rich
F:236-247/Region: glycine-rich
F:267-291/Region: histidine/proline-rich
F:316-383/Domain: POU domain homology <POU>
F:402-458/Domain: homeobox homology <HOX>

Query Match 16.0%; Score 81.5; DB 1; Length 495;
Best Local Similarity 26.1%; Pred. No. 0.5;
Matches 24; Conservative 7; Mismatches 36; Indels 25; Gaps 2;

QY 19 PAGFRHGGVQVHVKEFEVDVTSRAGANHHHH-----GHHGGHG- 62
DB 237 PGGGGGGAGGAGSLVHPLVGRDTPELAHHHHHHHHHPPHPPHQAQPPHHGGGA 296
QY 63 -----FVYRTRVEDINTCTGEVHER 85
DB 297 GPCLNSHDPHSDPTSDLEQFAKQKQRR 328

RESULT 10

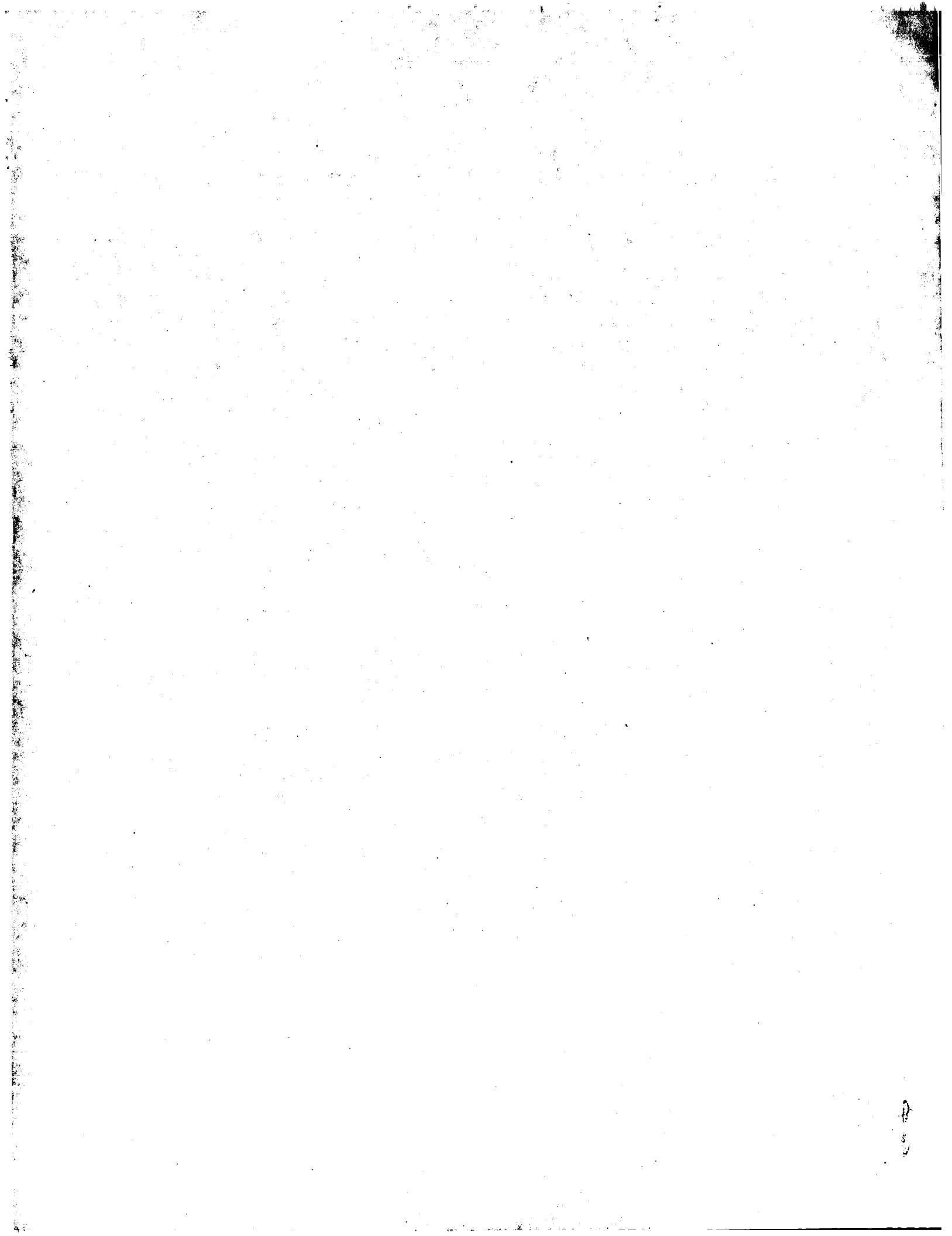
T16437
hypothetical protein F53A9.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16437
R:Miller, N.
submitted to the EMBL Data Library, March 1995
A:Description: The sequence of C. elegans cosmid F53A9.
A:Reference number: Z18513
A:Accession: T16437
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <MIL>
A:Cross-references: EMBL:U23523; NID:g746551; PID:g746557; PIDN:AAC46561.1; CESP:F53A
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F53A9.6

Query Match 15.9%; Score 81; DB 2; Length 86;
Best Local Similarity 29.5%; Pred. No. 0.089;
Matches 23; Conservative 5; Mismatches 24; Indels 26; Gaps 3;

QY 1 MAYTQYDYCSEEVRSVAPAGFRHGG-----GVQHVYKKEFEVDVTS 45

RESD01 13
T00338
hypothetical protein KIAA0570 - human

Search completed: June 6, 2003, 12:53:25
Job time : 26 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2003, 11:41:43 ; Search time 125 Seconds
(without alignments)
1077.357 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQEVDCYSEEVSRVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10090035/runat_06062003_105507_11116/app_query.fasta_1.526
-DB=PublishedApplications_NA -QFMT=fastap -SURFIX=rpnp -MINMATCH=0.1
-LOOCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10090035 -CGN_1_176 @runat_06062003_105507_11116
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	574	9 US-10-090-035-3	Sequence 3, Appli
2	510	100.0	577	9 US-10-090-035-5	Sequence 5, Appli
3	510	100.0	676	9 US-10-090-035-1	Sequence 1, Appli
4	507	99.4	529	9 US-10-090-035-9	Sequence 9, Appli

5	499.5	97.9	524	9	US-10-090-035-17	Sequence 17, Appli
6	499.5	97.9	580	9	US-10-090-035-7	Sequence 7, Appli
7	295	57.8	591	9	US-10-090-035-15	Sequence 15, Appli
8	277.5	54.4	436	9	US-10-090-035-21	Sequence 21, Appli
9	258	50.6	584	9	US-10-090-035-19	Sequence 19, Appli
10	258	50.6	584	9	US-10-090-035-13	Sequence 23, Appli
11	251	49.2	348	9	US-10-090-035-23	Sequence 13, Appli
12	213	41.8	237	10	US-09-923-876-2788	Sequence 2788, Ap
13	92.5	18.1	328	10	US-09-924-035A-17	Sequence 17, Appli
14	92.5	18.1	591	9	US-09-938-842A-1858	Sequence 1858, Ap
15	88.5	17.4	546	10	US-09-728-445-817	Sequence 817, App
16	83	16.3	3870	9	US-09-712-363-113	Sequence 113, App
17	82.5	16.2	3966	9	US-10-198-846-13254	Sequence 13254, A
18	80	15.7	10062	10	US-09-888-615-5	Sequence 5, Appli
19	78.5	15.4	1602	10	US-09-833-790-417	Sequence 417, App
20	78	15.3	1059	9	US-10-232-563-4	Sequence 4, Appli
21	78	15.3	1189	10	US-09-900-237-5	Sequence 5, Appli
22	78	15.3	2681	9	US-10-232-563-1	Sequence 1, Appli
23	77	15.1	1080	9	US-10-232-563-5	Sequence 5, Appli
24	77	15.1	6381	10	US-09-969-347-216	Sequence 216, App
25	76	14.9	1203	9	US-09-879-312-1	Sequence 1, Appli
26	76	14.9	2627	9	US-10-037-270-434	Sequence 434, App
27	75.5	14.8	491	9	US-09-854-133-343	Sequence 343, App
28	75.5	14.8	491	9	US-10-291-737-3	Sequence 3, Appli
29	75.5	14.8	25309	9	US-09-738-626-2739	Sequence 2739, Ap
30	75	14.7	8979	9	US-10-103-313-83	Sequence 83, Appli
31	75	14.7	3309400	9	US-10-153-668-95	Sequence 95, Appli
32	74.5	14.6	1458	9	US-09-998-660-1	Sequence 1, Appli
33	74.5	14.6	1519	9	US-09-864-761-6743	Sequence 325, App
34	74	14.5	464	10	US-10-062-254-325	Sequence 9679, Ap
35	74	14.5	558	9	US-09-764-891-9679	Sequence 143, App
36	74	14.5	593	10	US-09-984-130-143	Sequence 1494, Ap
37	74	14.5	910	12	US-10-052-798-7	Sequence 6, Appli
38	74	14.5	32082	9	US-10-278-173-15	Sequence 7, Appli
39	73.5	14.4	279	9	US-09-938-842A-457	Sequence 457, App
40	73.5	14.4	858	9	US-09-918-995-6480	Sequence 6480, Ap
41	73.5	14.4	930	12		
42	73.5	14.4	939	12		
43	73.5	14.4	1084	9		
44	73.5	14.4	2331	9		
45	73	14.3	472	9		

ALIGNMENTS

RESULT 1
US-10-090-035-3
; Sequence 3, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090.035
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/27/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(374)
US-10-090-035-3

Alignment Scores:
Pred No.: 1.3e-58
Score: 510.00
Percent Similarity: 100.00%
Length: 574
Matches: 93
Conservative: 0

```
Best-Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-3 (1-574)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 96 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGTGGCCCGGCC 155
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 156 GGCTTCGGCCGACGGCGCGCGCTCCAGCAGCAGCTCGTCAAGAGAAGTTCGAGGAG 215
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 216 GTCGACAGGTCTCAGCGCGCGCGCCAAACCACCACCACCATGGTCACCACGGCGC 275
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 276 CACGGCTTCGTGGTGGCGGAGCCAGGGTCGAAGAGGACATCAACACCTGCACCGCGC 335
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 336 GTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGGCTAAC 374

RESULT 2
US-10-090-035-5
; Sequence 5, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(377)
US-10-090-035-5
Alignment Scores:
Pred. No.: 1,31e-58 Length: 577
Score: 510.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-5 (1-577)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 99 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGTGGCCCGGCC 159
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 159 GGCTTCGGCCGACGGCGCGCGCTCCAGCAGCAGCTCGTCAAGAGAAGTTCGAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 219 GTCGACAGGTCTCAGCGCGCGCGCCAAACCACCACCACCATGGTCACCACGGCGC 278
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80

US-10-090-035-4 (1-93) x US-10-090-035-1 (1-676)

QY 1 MetAlaTyrTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
DB 89 ATGGCTTACTACAGGAGGTGGACTACTCTCGGAGGAGGTGAGTCTGGTGGCCCGGCC 148
QY 21 GlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGluGlu 40
DB 149 GGCTTCGGCCGACGGCGCGCGCTCCAGCAGCAGCTCGTCAAGAGAAGTTCGAGGAG 208
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGlyGly 60
DB 209 GTCGACAGGTATCAGCGCGCGCGCCAAACCACCACCACCATGGTCACCACGGCGC 268
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspIleAsnThrCysThrGlyGlu 80
DB 269 CACGGCTTCGTGGTGGCGGAGACCGAGGTGCGAGGAGGACATCAACACCTGCACCGCGC 328
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
DB 329 GTCCACGAGCGCAGGAGGAGCTTCCTCGCCAGGGCTAAC 367

RESULT 4
US-10-090-035-9
; Sequence 9, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
```


;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9

;; LENGTH: 529
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (53)...(331)
US-10-090-035-9

Alignment Scores:
Pred. No.: 2,93e-58 Length: 529
Score: 507.00 Matches: 92
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.92% Mismatches: 0
Query Match: 99.41% Indels: 0
DB: Gaps: 0

US-10-090-035-4 (1-93) x US-10-090-035-9 (1-529)

QY 1 MetAlaTyrTyTGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 53 ATGGCTTACTACCGAGAGGTGACTACTCGCGAGGAGGTGAGTGGTGGCCCGGCC 112
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValLysGluLysPheGluGlu 40
|||||
Db 113 GGCTTCGGCGCCGACGCGCGCGGCTCCAGGAGGAGTTCGAGGAG 172
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisHis 60
|||||
Db 173 GTCGACACGGTTCGACGCGCGCGGCTCCAGGAGGAGTTCGAGGAG 172
QY 61 HisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGlyGlu 80
|||||
Db 233 CACGGCTTCGTGGTGGCGGACACCGGTCGAGGAGGAGGACATCAACACCTGCACCGCGGAG 292
QY 81 ValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 293 GTCCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 331

RESULT 5

US-10-090-035-17
;; Sequence 17, Application US/100900035
;; Patent No. US20020170089A1
;; GENERAL INFORMATION:
;; APPLICANT: Simmons, Carl R.
;; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
;; FILE REFERENCE: 35718/242990
;; CURRENT APPLICATION NUMBER: US/10/090,035
;; PRIOR FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: 60/272,227
;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17

;; LENGTH: 524
;; TYPE: DNA
;; ORGANISM: Triticum aestivum
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (57)...(338)
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(524)
;; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-17

Alignment Scores:
Pred. No.: 2,85e-57 Length: 524
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0

Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: Gaps: 1

US-10-090-035-4 (1-93) x US-10-090-035-17 (1-524)

QY 1 MetAlaTyrTyTGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 57 ATGGCTTACTACCGAGGAGTGGACTACTCGCGAGGAGGTGAGTGGTGGCCCGGCC 116
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValLysGluLysPheGluGlu 40
|||||
Db 117 GGCTTCGGCGCCGACGCGCGGCTCCAGGAGGAGTTCGAGGAG 176
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisHis 59
|||||
Db 177 GTCGACACGGTTCACGCGCGCGGCTCCAGGAGGAGTTCGAGGAG 236
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGly 79
|||||
Db 237 GGCCACGGCTTCGTGGTGGCGGAGACGAGGTCGAGGAGGACATCAACACCTGCACCGCG 296
QY 80 GluValHisGluArgArgGluSerPheLeuAlaArgAlaAsn 93
Db 297 GAGGTCACGAGCGCAGGAGAGCTTCCTCGCCAGGCGCTAAC 338

RESULT 6

US-10-090-035-7
;; Sequence 7, Application US/100900035
;; Patent No. US20020170089A1
;; GENERAL INFORMATION:
;; APPLICANT: Simmons, Carl R.
;; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
;; FILE REFERENCE: 35718/242990
;; CURRENT APPLICATION NUMBER: US/10/090,035
;; CURRENT FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: 60/272,227
;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7

;; LENGTH: 580
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (99)...(380)
US-10-090-035-7

Alignment Scores:

Pred. No.: 3,24e-57 Length: 580
Score: 499.50 Matches: 93
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 97.94% Indels: 1
DB: Gaps: 1

US-10-090-035-4 (1-93) x US-10-090-035-7 (1-580)

QY 1 MetAlaTyrTyTGlnGluValAspTyrCysSerGluGluValArgSerValAlaProAla 20
|||||
Db 99 ATGGCTTACTACCGAGGAGTGGACTACTCGCGAGGAGGTGAGTGGTGGCCCGGCC 158
QY 21 GlyPheGlyArgHisGlyGlyGlyValGlnGlnHisValLysGluLysPheGluGlu 40
|||||
Db 159 GGCTTCGGCGCCGACGCGCGGCTCCAGGAGGAGTTCGAGGAG 218
QY 41 ValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisHisHis 59
|||||
Db 219 GTCGACACGGTTCACGCGCGCGGCTCCAGGAGGAGTTCGAGGAG 278
QY 60 GlyHisGlyPheValValArgGluThrArgValGluGluAspPheAsnThrCysThrGly 79

Db 279 GCCACGGCTTGGTGGCGAGACAGGGTCGAGGAGGACATCAACACCTGCACCGGC 338
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 339 GAGTCCAGGAGCGGAGGAGAGCTTCTCGCCAGGGCTAAC 380
RESULT 7
US-10-090-035-15
; Sequence 15, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(333)
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-15
Alignment Scores:
Pred. No.: 4e-30 Length: 591
Score: 295.00 Matches: 57
Percent Similarity: 75.53% Conservatives: 14
Best Local Similarity: 60.64% Mismatches: 19
Query Match: 57.84% Indels: 4
Gaps: 2
US-10-090-035-4 (1-93) x US-10-090-035-15 (1-591)
QY 1 MetaLactyTyrGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
Db 61 ATGGCTCATTACAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGTGACCCGACC 120
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
Db 121 GCGGCTTCTCGCGCGCGCGGTGCAGCAGCAGCAGCTGCTCAAGGAGAGCGTTCCAG 180
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisGly 59
Db 181 GAGATCGAC-----AGTCCGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231
QY 60 GlyHisGlyPheValArgGluThrArgValGluGluAspLeuAsnThrCysThrGly 79
Db 232 AACGACTACCTGATGTTGCGGAGACCAAGGTGAGGAGGAGCTTCACACCTGCACCGGC 291
QY 80 GluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
Db 292 GAGTCCGCGAGCGCAAGCAGAGCTTCTGCTCAAGTCCGAC 333
RESULT 8
US-10-090-035-21
; Sequence 21, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990

; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(326)
US-10-090-035-21
Alignment Scores:
Pred. No.: 5.68e-28 Length: 436
Score: 277.50 Matches: 60
Percent Similarity: 73.47% Conservatives: 12
Best Local Similarity: 61.22% Mismatches: 13
Query Match: 54.41% Indels: 13
Gaps: 6
US-10-090-035-4 (1-93) x US-10-090-035-21 (1-436)
QY 1 MetaLactyTyrGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
Db 54 ATGGCGCACTTCCAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGGCGGTGGGCAACCCG 113
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPhe--- 38
Db 114 GCC-----CGCCGCGCGCGCGGTGCAGGAGCAGCATCGTCAAGGAGACGTTCTGTG 164
QY 39 GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisGlyHisHis 58
Db 165 CAGGAGTTCGACACCTCCGCGCGC-----CGCCACGCTCACCGGTCCAC 212
QY 59 Gly-----GlyHisGlyPheValValArgGluThrArgValGluGluAspLeuAsn 75
Db 213 GGCCGCGCGCTCTGGTGCAC---TTCGAGTGGCGGAGAGCAGGCTCGAGGAGACTTCAAC 269
QY 76 ThrCysThrGlyGluValHisGlyGlyValGlnGlnHisGlySerPheLeuAlaArgAlaAsn 93
Db 270 ACCCGCAGCGGGGAGTTCACGAGCGCAAGGAGAACTTCGTCGTCAGGCGCGCAT 323
RESULT 9
US-10-090-035-19
; Sequence 19, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-19
Alignment Scores:

Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-4 (1-93) x US-10-090-035-19 (1-584)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 46 ATGGCGCACTTCCAGGAGTGGACTACTGCTCGAGGAGGTGAGGCGGTGGCTACCCG 105
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 106 GCC-----CGCCGCGCTGCGGCGCGTGCAGGAGCACATCGTCAAGGAGCATTC 156
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHis 57
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 157 GTGAGGAGTTCGACACCGCGCGCG-----CGCCAYGGTCCACCGGTAC 204
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluAspIle 74
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 205 CAGCGCGCGTCTCGGTCTAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGGACATC 261
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 262 AACACCGCGCAGGAGTTCACGACGCAAGGAAATTCCTCTCCAGGCGCAT 318
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 10

```
US-10-090-035-23
; Sequence 23, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(321)
; NAME/KEY: misc_feature
; LOCATION: (1)...(584)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-23
```

Alignment Scores:
Pred. No.: 3.13e-25 Length: 584
Score: 258.00 Matches: 58
Percent Similarity: 72.73% Conservative: 14
Best Local Similarity: 58.59% Mismatches: 13
Query Match: 50.59% Indels: 14
DB: 9 Gaps: 7

US-10-090-035-4 (1-93) x US-10-090-035-23 (1-584)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAla---Pro 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 46 ATGGCGCACTTCCAGGAGTGGACTACTGCTCGAGGAGGTGAGGCGGTGGCTACCCG 105
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 20 AlaGlyPheGlyArgHisGly---GlyGlyValGlnGlnHisValValLysGluLysPhe 38
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 106 GCC-----CGCCGCGCTGCGGCGCGTGCAGGAGCACATCGTCAAGGAGCATTC 156
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
QY 39 ---GluGluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHis 57
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 157 GTGAGGAGTTCGACACCGCGCGCG-----CGCCAYGGTCCACCGGTAC 204
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 58 HisGly-----GlyHisGlyPheValValArgGluThrArgValGluAspIle 74
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 205 CAGCGCGCGTCTCGGTCTAC---TTCGAGGTGCGGAGAGCAAGCTSGARGAGGACATC 261
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 75 AsnThrCysThrGlyGluValHisGluArgGluSerPheLeuAlaArgAlaAsn 93
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 262 AACACCGCGCAGGAGTTCACGACGCAAGGAAATTCCTCTCCAGGCGCAT 318
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 11

```
US-10-090-035-13
; Sequence 13, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(348)
; NAME/KEY: misc_feature
; LOCATION: (1)...(348)
; OTHER INFORMATION: n = A,T,C or G
US-10-090-035-13
```

Alignment Scores:
Pred. No.: 1.39e-24 Length: 348
Score: 251.00 Matches: 51
Percent Similarity: 70.59% Conservative: 9
Best Local Similarity: 60.00% Mismatches: 21
Query Match: 49.22% Indels: 4
DB: 9 Gaps: 2

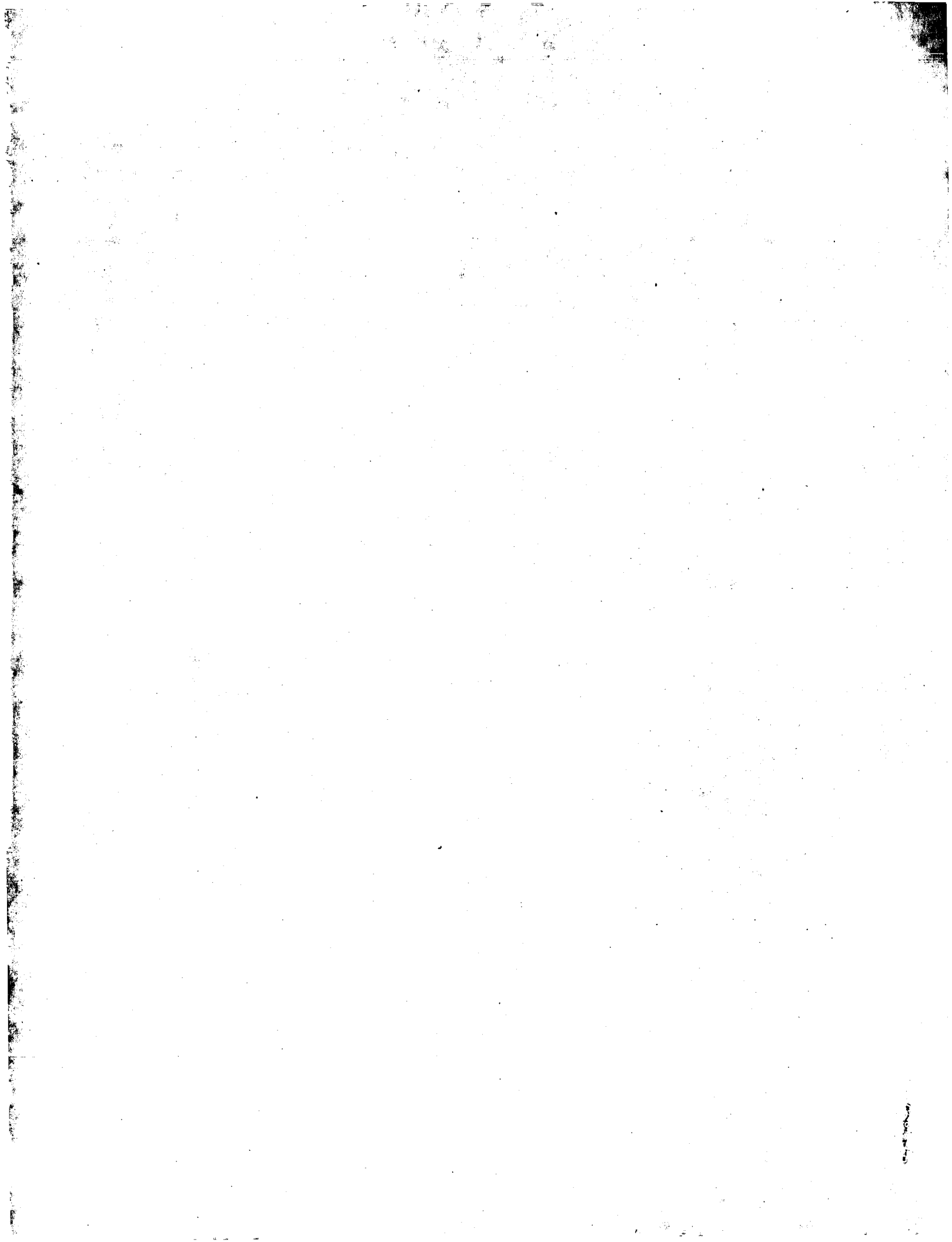
US-10-090-035-4 (1-93) x US-10-090-035-13 (1-348)

```
QY 1 MetAlaTyrTrpGlnGluValAspTyrCysSerGluGluValArgSerValAlaPro--- 19
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 52 ATGGCTCACTTACCAGGAGTGGACTACTGCTCGGAGGAGGTGAGGTGCGGCCACCC 111
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 20 AlaGlyPheGlyArgHisGlyGlyValGlnGlnHisValValLysGluLysPheGlu 39
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 112 GCGCGTCTCTCGCGCGCGCGCGTGCAGCAGCAGCAGCGTCTCAAGGAGAGCTTCCAG 171
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 40 GluValAspThrValSerArgAlaGlyAlaAsnHisHisHisHisHisHisHisGly 59
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 172 GAGATCGACANG-----TCCGGCTCCGCGCGGCAGCAGCAGCAGCAGCAGCAGC 222
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 GlyHisGlyPheValValArgGluThrArgValGluAspIleAsnThrCysThrGly 79
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 223 AAGCACTACTTATGTGCGGAGACCAAGGTGAGGAGACTTTTACACCTCACCGGC 282
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 80 GluValHisGluArg 84
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 283 GAGTTTCGCGAGCGC 297
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

RESULT 12

US-09-923-876-2788
; Sequence 2788, Application US/09923876

Search completed: June 16, 2003, 13:51:21
Job time : 130 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:52:33 ; Search time 29 Seconds
(without alignments)
331.081 Million cell updates/sec

Title: US-10-090-035-4

Perfect score: 510

Sequence: 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQVQVYKFEEDTVSRAGANHHHHHHGG 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	93	9	US-10-090-035-2
2	510	100.0	93	9	US-10-090-035-4
3	510	100.0	93	9	US-10-090-035-6
4	507	99.4	93	9	US-10-090-035-10
5	499.5	97.9	94	9	US-10-090-035-8
6	499.5	97.9	94	9	US-10-090-035-18
7	295	57.8	91	9	US-10-090-035-16
8	277.5	54.4	91	9	US-10-090-035-22
9	273	53.5	92	9	US-10-090-035-20
10	258	50.6	92	9	US-10-090-035-24
11	245	48.0	99	9	US-10-090-035-14
12	80	15.7	359	9	US-10-232-563-7
13	80	15.7	359	10	US-09-888-615-64
14	78	15.3	352	9	US-10-232-563-2
15	77	15.1	359	9	US-10-232-563-6
16	76	14.9	400	9	US-09-879-312-2
17	73.5	14.4	285	9	US-10-278-173-16
18	73.5	14.4	309	12	US-10-052-798-9
19	73.5	14.4	312	12	US-10-052-798-10

20	73	14.3	19	9	US-09-876-904A-627	Sequence 627, App
21	73	14.3	633	10	US-09-824-735-3	Sequence 3, Appl1
22	73	14.3	633	10	US-09-801-368-338	Sequence 338, App
23	71.5	14.0	310	12	US-10-052-798-11	Sequence 11, Appl
24	71.5	14.0	354	10	US-09-925-302-501	Sequence 501, App
25	71.5	14.0	532	10	US-09-833-790-428	Sequence 428, App
26	69.5	13.6	315	9	US-10-232-563-11	Sequence 11, Appl
27	69.5	13.6	623	9	US-10-108-605-125	Sequence 125, App
28	69.5	13.6	623	9	US-10-108-605-129	Sequence 129, App
29	69.5	13.6	726	9	US-09-932-257A-19	Sequence 19, Appl
30	69	13.5	257	9	US-09-738-626-6135	Sequence 6135, Ap
31	68.5	13.4	20	9	US-09-051-013-3	Sequence 3, Appl1
32	68	13.3	139	10	US-09-813-820-8	Sequence 8, Appl1
33	67.5	13.2	118	10	US-09-205-658-120	Sequence 120, App
34	67.5	13.2	530	9	US-10-044-692-317	Sequence 317, App
35	67.5	13.2	530	9	US-10-044-539-317	Sequence 317, App
36	67.5	13.2	567	9	US-10-270-333-126	Sequence 126, App
37	67.5	13.2	605	9	US-09-741-233A-2	Sequence 2, Appl1
38	67.5	13.2	1207	9	US-10-108-605-71	Sequence 71, Appl
39	67	13.1	446	10	US-09-853-386-69	Sequence 69, Appl
40	67	13.1	507	9	US-09-795-927-10	Sequence 10, Appl
41	67	13.1	507	10	US-09-738-897-2	Sequence 201, App
42	67	13.1	608	9	US-09-975-719-201	Sequence 201, App
43	67	13.1	639	9	US-09-975-719-200	Sequence 199, App
44	67	13.1	643	9	US-09-975-719-199	Sequence 70, Appl
45	67	13.1	643	10	US-09-853-386-70	

ALIGNMENTS

RESULT 1
US-10-090-035-2
; Sequence 2, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Zea mays
US-10-090-035-2

Query Match Best Local Similarity 100.0%; Score 510; DB 9; Length 93;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQVQVYKFEEDTVSRAGANHHHHHHGG 60
Db 1 MAYTQVDYCSSEVRSVAPAGFGRHGGVQVQVYKFEEDTVSRAGANHHHHHHHHGG 60

Qy 61 HGFVVRTRVEEDINTCTGVEHRRSFARAN 93
Db 61 HGFVVRTRVEEDINTCTGVEHRRSFARAN 93

RESULT 2
US-10-090-035-4
; Sequence 4, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; Proteins and Uses Thereof

```

; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-090-035-4

Query Match          100.0%; Score 510; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 3
US-10-090-035-6
; Sequence 6, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-090-035-6

Query Match          100.0%; Score 510; DB 9; Length 93;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 4
US-10-090-035-10
; Sequence 10, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; US-10-090-035-10

Query Match          99.4%; Score 507; DB 9; Length 93;
Best Local Similarity 98.9%; Pred. No. 1e-44;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Db 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93

RESULT 5
US-10-090-035-8
; Sequence 8, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-090-035-8

Query Match          97.9%; Score 499.5; DB 9; Length 94;
Best Local Similarity 98.9%; Pred. No. 5.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 59
Db 1 MAYTQEDYCSSEVRSVAPAGFGRHGGVQOHVVKKEFEVDTVSRAGANHHHHHHGG 60
Qy 60 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 93
Db 61 HGFVVRTRVEEDINTCTGEVHERRESFLARAN 94

RESULT 6
US-10-090-035-18
; Sequence 18, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 94
; TYPE: PRT
; US-10-090-035-18
```



```

49  A-----NHHHHGHHGGHG 62
      |
175  ADDMGAGHHHGAHHTAHHHHSANHHHHHHHHGGSG 210

```

RESULT 13

```

US-09-888-615-64
; Sequence 64, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENPEPEL, SEAN
; APPLICANT: CHARVICKZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 3353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1891)
; OTHER INFORMATION: Any amino acid
US-09-888-615-64

```

	Query Match	15.7%	Score 80;	DB 10;	Length 3353;	
	Best Local Similarity	33.3%;	Pred. No. 14;			
	Matches	15;	Conservative	9;	Mismatches	19; Indels
						2; Gaps
QY	48	GANHNNHHGHGGH--GFVVRETRVREDNTCTGVEHRRESPLA	90			
		: :	:	:	:	:
Dh	746	GPOHHNNHHNNHHNHHCHMMDDMLSLADDVSCSSQSVAKSERKMA	790			

RESULT 14

```

US-10-232-563-2
, Sequence 2, Application US/10232563
, Publication No. US20030087394A1
, GENERAL INFORMATION:
, APPLICANT: Sharma, Arun
, TITLE OF INVENTION: INSULIN RELATED TRANSCRIPTION FACTOR AND
, TITLE OF INVENTION: USES THEREOF
, FILE REFERENCE: 10276-072001
, CURRENT APPLICATION NUMBER: US/10/232,563
, CURRENT FILING DATE: 2002-08-30
, PRIOR APPLICATION NUMBER: US 60/316,453
, PRIOR FILING DATE: 2001-08-31
, NUMBER OF SEQ ID NOS: 22
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 2
, LENGTH: 352
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-232-563-2

```

Query Match	15.3%	Score 78;	DB 9;	Length 352;
Best Local Similarity	29.9%;	Pred. No. 2.1;		
Matches	23;	Conservative	7;	Mismatches 27;
				Indels 20;
				Gaps 5;
QY	21	GFGRGGVGQVHVKEEFDVTSRAGANHHHHHHG--GHG-	:-:	
				-----FVVRTRVEED 73
DB	181	GAGHHGA--HAAHH-----HHAHHHHHHHHGGAGGCGAGGHVRLERTSDD		229
QY	74	--INTCTGEVHERRESF	88	

58 HG---GHGFVRETRVEEDINTCTGEVHERRESFLARAN.93
.:||| || | |||:::||| ||| ||| :| ::|:
54 HCRGSGH-FFEVRESKLEEDINTRTGETHERKGNFSSKAD.91

RESULT 11

```

US-10-090-035-14
; Sequence 14, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 02/272,227
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(99)
; OTHER INFORMATION: xaa = Any Amino Acid
US-10-090-035-14

```

	Query Match	48.08;	Score 245;	DB 9;	Length 99;
	Best Local Similarity	58.88;	Fred. No. 5.4e-18;		
	Matches 50;	Conservative	Mismatches 22;	Indels 4;	Gaps 2;
Qy	1 MAYYQEYDYCSEEVRSVAP-AGGRHGGGVQOHHVVEKFEEDTVSRAGANHHHHHHCHHG	59			
	1 MAHYQEYDYCSEEVRSVPTGGFLRGGVQOHHVVKETFEIDX----	57			
Db	60 GHGFVVRVRETRVEEDINTCTGEVHER	84			
Qy	58 NDYXWVRKTKXEEDNTCTGEFR	82			

REFSUT 12

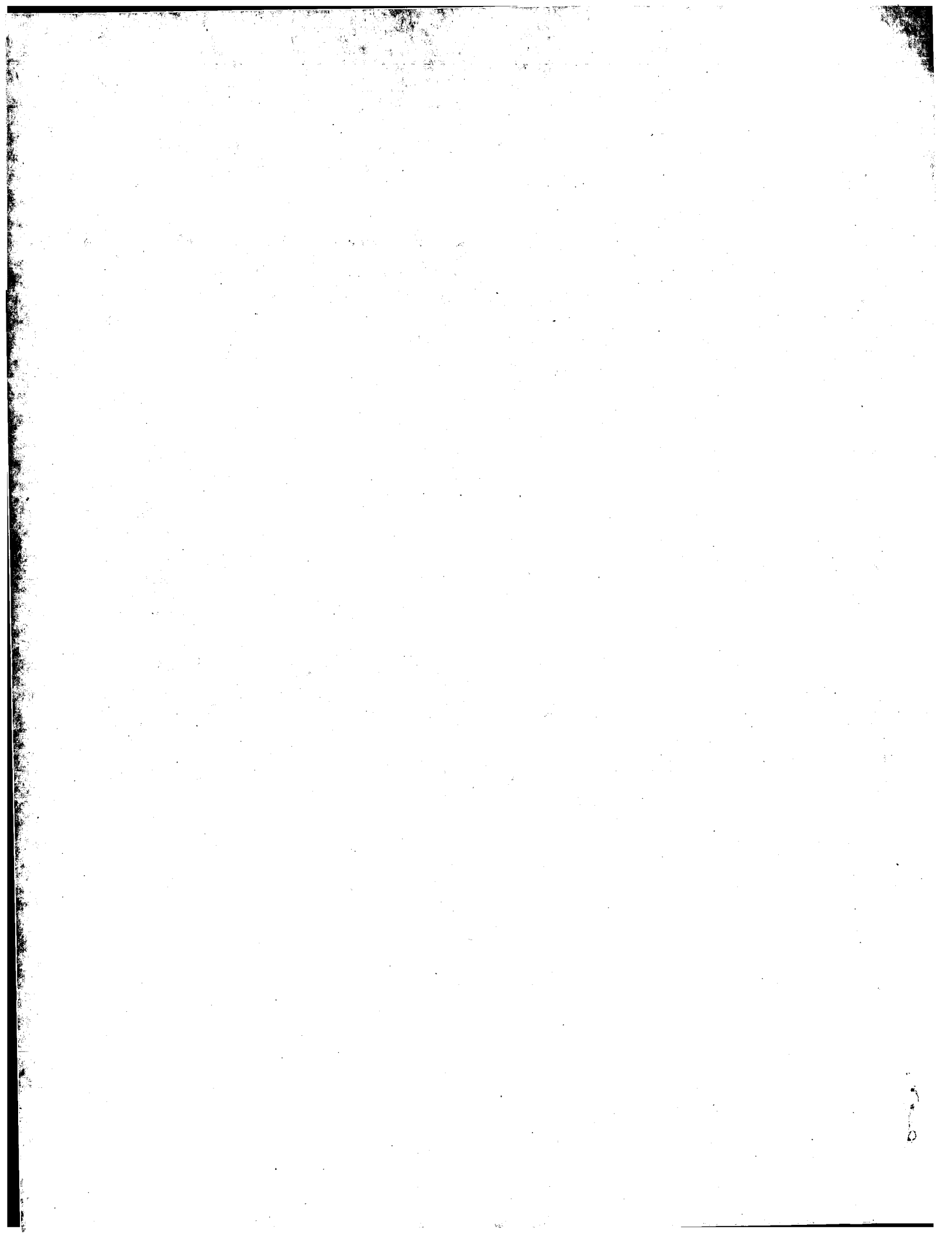
```

US-10-232-563-7
; Sequence 7, Application US/10232563
; Publication NO. US20030087394A1
; GENERAL INFORMATION:
; APPLICANT: Sharma, Arun
; TITLE OF INVENTION: INSULIN RELATED TRANSCRIPTION FACTOR AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 10276-072001
; CURRENT APPLICATION NUMBER: US/10/232,563
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/316,453
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-232-563-7

```

```
Query Match      15.7%   Score 80; DB 9; Length 359;
Best Local Similarity 26.0%; Pred.No. 1.4;
Matches 25; Conservative 7; Mismatches 28; Indels 36; Gaps 4;
```

Search completed: June 6, 2003, 13:02:10
Job time : 29 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:49:03 ; Search time 18 Seconds
(Without alignments)
152.018 Million cell updates/sec

Title: US-10-090-035-4

Perfect score: 510

Sequence: 1 MAYOQVDYCSSEVRVAPV.....INTCTGVEHRESFLARAN 93

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	15.5	414	5	PCT-US92-06840-2
2	77	15.1	763	2	US-08-677-862-2
3	77	15.1	763	2	US-09-252-571-2
4	77	15.1	763	3	US-09-434-065-2
5	77	15.1	763	4	US-08-789-275-4
6	77	15.1	763	4	US-08-789-275-5
7	76	14.9	400	4	US-09-086-010-2
8	73.5	14.4	40	2	US-08-273-146-39
9	73.5	14.4	309	4	US-09-079-029-9
10	73.5	14.4	312	4	US-09-079-029-10
11	73	14.3	617	1	US-08-137-614A-26
12	73	14.3	633	4	US-08-557-006C-43
13	73	14.3	637	3	US-08-072-064-1
14	73	14.3	637	3	US-08-072-064-4
15	73	14.3	637	3	US-08-072-064-6
16	73	14.3	637	3	US-08-072-064-8
17	73	14.3	637	5	PCT-US92-08558-1
18	72.5	14.2	79	2	US-08-448-418-97
19	71.5	14.0	310	4	US-09-079-029-11
20	70.5	13.8	284	3	US-09-184-658-40
21	70	13.7	398	4	US-09-461-474-17
22	69.5	13.6	726	4	US-09-126-980-2
23	69.5	13.6	726	4	US-09-476-482-2
24	68.5	13.4	281	4	US-09-517-605-6
25	68.5	13.4	281	4	US-09-423-439-44
26	68	13.3	139	4	US-08-856-253-8
27	67.5	13.2	379	4	US-09-457-040B-36

28	67.5	13.2	530	4	US-08-974-549A-603	Sequence 603, App
29	67	13.1	608	4	US-09-199-637A-201	Sequence 201, App
30	67	13.1	639	4	US-09-199-637A-200	Sequence 200, App
31	67	13.1	643	4	US-09-199-637A-199	Sequence 199, App
32	67	13.1	645	4	US-09-199-637A-253	Sequence 253, App
33	66.5	13.0	449	2	US-08-927-394-2	Sequence 2, Appli
34	66	12.9	159	3	US-08-991-890-4	Sequence 4, Appli
35	66	12.9	349	4	US-03-011-769A-47	Sequence 47, Appl
36	66	12.9	349	4	US-09-011-769A-60	Sequence 60, Appl
37	66	12.9	349	4	US-09-011-769A-64	Sequence 64, Appl
38	66	12.9	379	1	US-08-279-270A-1	Sequence 1, Appli
39	65.5	12.8	474	2	US-08-650-000-4	Sequence 4, Appli
40	65.5	12.8	474	4	US-09-042-785A-8	Sequence 8, Appli
41	65.5	12.8	474	6	5395760-4	Patent No. 5395760
42	65.5	12.8	705	2	US-08-770-761A-7	Sequence 7, Appli
43	64.5	12.6	2353	4	US-08-984-709A-50	Sequence 50, Appl
44	64	12.5	60	1	US-08-255-457-1	Sequence 1, Appli
45	64	12.5	60	2	US-09-115-032-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
PCT-US92-06840-2
; Sequence 2, Application PC/TUS9206840
; GENERAL INFORMATION:
; APPLICANT: Shi, Yang
; APPLICANT: Seto, Edward
; APPLICANT: Shenk, Thomas
; TITLE OF INVENTION: Y1 TRANSCRIPTION FACTOR AND METHODS OF
; TITLE OF INVENTION: ISOLATING SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas - 7th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06840
; FILING DATE: 19920814
; CLASSIFICATION:
; PRIOR APPLICATION: AU 1805
; APPLICATION DATA: US 07/746,485
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Dennis, Manette
; REGISTRATION NUMBER: 30,623
; REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-06840-2

Query Match 15.5% Score 79; DB 5; Length 414;
Best Local Similarity 29.2%; Pred. No. 0.088;
Matches 21; Conservative 6; Mismatches 25; Indels 20; Gaps 3;

```

QY 21 GFRHGGVQVHVKEEEDTVSRAGANHHHHHGGH-GFVVRTRVEEDINTCTG 79
Db 55 GGDHGGG-----GGHGHAGHHHHHHHHHPMIALQPLVTD-----PT. 95
QY 80 EVHERRESFLAR 91
Db 96 QVHHQEVILQ 107

RESULT 2
US-08-677-862-2
; Sequence 2, Application US/08677862
; Patent No. 5874230
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,862
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,862
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: T96-005/A63613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 781-1989
; TELEFAX: 415 398-3249
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-677-862-2

Query Match 15.1%; Score 77; DB 2; Length 763;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;

QY 11 SEEVRSVAPAGFGRHGG-----GVOQHVKKEEEDTVSRA-----GANHHHHHG 56
Db 555 SPQVRQFPAPGLWSGTEAPQVTVEHPVQETTFHVAPOQNALHHHGGSSHHHHHHH 614
QY 57 HHGGHG 62
Db 615 HHHHHG 620

RESULT 3
US-09-434-065-2
; Sequence 2, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/434,065
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; US-09-434-065-2

Query Match 15.1%; Score 77; DB 2; Length 763;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;

QY 11 SEEVRSVAPAGFGRHGG-----GVOQHVKKEEEDTVSRA-----GANHHHHHG 56
Db 555 SPQVRQFPAPGLWSGTEAPQVTVEHPVQETTFHVAPOQNALHHHGGSSHHHHHHH 614
QY 57 HHGGHG 62
Db 615 HHHHHG 620

RESULT 4
US-09-434-065-2
; Sequence 2, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; APPLICANT: ROTHE, Mike
; TITLE OF INVENTION: TRAF2-Associated Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/434,065
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; US-09-434-065-2

Query Match 15.1%; Score 77; DB 2; Length 763;
Best Local Similarity 33.3%; Pred. No. 0.33;
Matches 22; Conservative 4; Mismatches 26; Indels 14; Gaps 2;

QY 11 SEEVRSVAPAGFGRHGG-----GVOQHVKKEEEDTVSRA-----GANHHHHHG 56
Db 555 SPQVRQFPAPGLWSGTEAPQVTVEHPVQETTFHVAPOQNALHHHGGSSHHHHHHH 614
QY 57 HHGGHG 62
Db 615 HHHHHG 620
```

RESULT 6
US-08-789-275-5
; Sequence 5, Application US/08789275A
; Patent No. 6251664
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```


APPLICATION NUMBER: US/09/079,029
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P101R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5416
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-079-029-10

Query Match 14.4%; Score 73.5; DB 4; Length 312;
Best Local Similarity 32.9%; Pred. No. 0.29;
Matches 25; Conservative 6; Mismatches 14; Indels 31; Gaps 6;
QY 6 EVD-YGSEVRVAPAGFRH---GGVVOHVVKKEFEVDVSRAGANHHHHHHGHHG 61
Db 259 EADYCNRSDDSS-----GNHVFGGTYKL-----TVLGAAHHHHHHHHG----- 296
QY 62 GFVRETRV--BEDIN 75
Db 297 ----AAEQKLISEDLN 309

RESULT 11
US-08-137-614A-26
Sequence 26, Application US/08137614A
Patent No. 5487976
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Knipple, Douglas C.
APPLICANT: Henderson, Joseph E.
TITLE OF INVENTION: Gene Encoding An Insect
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/137,614A
FILING DATE: 15-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-137-614A-26

Query Match 14.3%; Score 73; DB 1; Length 617;

Best Local Similarity 27.8%; Pred. No. 0.78;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;
QY 17 VAPAGFRHGGVQHVVKKEFEVDVSRAGANHHHHHHHHG---GHGFVRET----- 68
Db 413 VPGGPGGPGGGV-----NVGVGMGMGPEHGHGHHHSHGHAPKQTVSNRP 462
QY 69 ----RVEDINT 76
Db 463 IGFSNIQNVGT 474

RESULT 12
US-08-557-006C-43
Sequence 43, Application US/08557006C
Patent No. 6258547
GENERAL INFORMATION:
APPLICANT: Beri, Rajinder K.
APPLICANT: Carling, David
APPLICANT: Forster, Robert A.
TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PHM37588/UST
CURRENT APPLICATION NUMBER: US/08/557,006C
CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 43
LENGTH: 633
TYPE: PRT
ORGANISM: Yeast
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(633)
OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match 14.3%; Score 73; DB 4; Length 633;
Best Local Similarity 76.9%; Pred. No. 0.81;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 50 NHHHHHHHHGHHG 62
Db 21 HHHHHHHHHHHG 33

RESULT 13
US-08-072-064-1
Sequence 1, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Mon Jun 16 14:55:57 2003

```

; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; ORGANISM: Drosophila melanogaster
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: III; polytene subregion 66F
; MAP POSITION: approximately map unit 26
;
US-08-072-064-1
;
Query Match      14.3%; Score 73; DB 3; Length 637;
Best Local Similarity 27.8%; Pred. No. 0.81;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;

QY 17 VAPAGFRGGGVQOHVVKKEFEVDTVSRAGANHHHHHHG---GHGFVVRET----- 68
Db 434 VPGGPGPGGGV-----NVGVGMGMPGHEGHGHAHSHGHPAKQTVSNRP 483
QY 69 ----RVEDINT 76
Db 484 IGFSNIQNVGT 495

RESULT 15
US-08-072-064-6
; Sequence 6, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
US-08-072-064-6
;
Query Match      14.3%; Score 73; DB 3; Length 637;
Best Local Similarity 27.8%; Pred. No. 0.81;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;

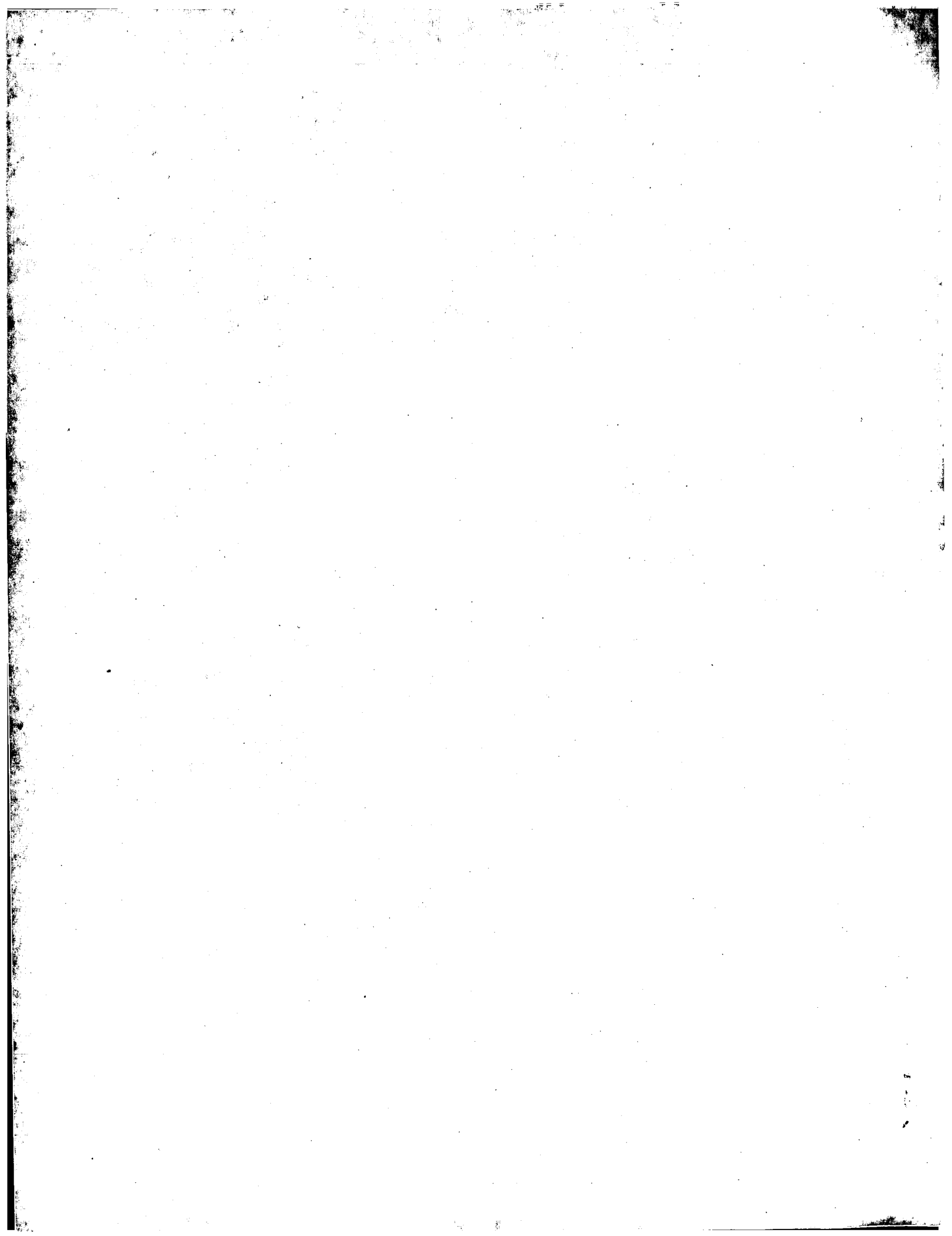
QY 17 VAPAGFRGGGVQOHVVKKEFEVDTVSRAGANHHHHHHG---GHGFVVRET----- 68
Db 434 VPGGPGPGGGV-----NVGVGMGMPGHEGHGHAHSHGHPAKQTVSNRP 483
QY 69 ----RVEDINT 76
Db 484 IGFSNIQNVGT 495

RESULT 14
US-08-072-064-4
; Sequence 4, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FFRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
;
US-08-072-064-4
;
Query Match      14.3%; Score 73; DB 3; Length 637;
Best Local Similarity 27.8%; Pred. No. 0.81;
Matches 20; Conservative 7; Mismatches 23; Indels 22; Gaps 3;

QY 17 VAPAGFRGGGVQOHVVKKEFEVDTVSRAGANHHHHHHG---GHGFVVRET----- 68
Db 434 VPGGPGPGGGV-----NVGVGMGMPGHEGHGHAHSHGHPAKQTVSNRP 483
QY 69 ----RVEDINT 76
Db 484 IGFSNIQNVGT 495
```

Oy 69 ----RVEDINT 76
Db 484 IGFSNIOQNVGT 495
::: |

Search completed: June 6, 2003, 12:54:08
Job time : 19 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 12:42:12 ; Search time 31.5 Seconds
(without alignments)
393.407 Million cell updates/sec

Title: US-10-090-035-4
Perfect score: 510
Sequence: 1 MAYQEVDYCEEVRSVAPA.....INTCTGEVHERRESFLARAN 93

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	A_Geneseq_101002.*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	16.5	441	21	Arabidopsis thalia
2	84	16.5	473	21	Arabidopsis thalia
3	84	16.5	488	21	Arabidopsis thalia
4	84	16.5	1561	22	Drosophila melanog
5	83	16.3	168	22	scfv lba antibody
6	81.5	16.0	271	23	Drosophila melanog
7	80.5	15.8	347	22	Arabidopsis thalia
8	80.5	15.8	384	21	Arabidopsis thalia
9	80.5	15.8	388	21	Arabidopsis thalia
10	80.5	15.8	388	21	Arabidopsis thalia

11	80.5	15.8	396	21	AAG13901
12	80.5	15.8	396	21	AAG43724
13	80.5	15.8	405	21	AAG49723
14	80	15.7	3353	23	AAU82706
15	79	15.5	87	22	AA111465
16	79	15.5	153	22	AA111466
17	79	15.5	153	22	AA111466
18	79	15.5	414	14	AA132020
19	79	15.5	414	19	AAW65406
20	78	15.3	1020	22	AA190775
21	77.5	15.2	537	22	AB171829
22	77.5	15.2	566	22	AB171829
23	77.5	15.2	1693	22	AB171829
24	77	15.1	509	22	AB171829
25	77	15.1	509	22	AB171829
26	77	15.1	509	22	AB171829
27	77	15.1	763	19	AAW41734
28	77	15.1	763	22	AAU02222
29	77	15.1	763	22	AAU02222
30	77	15.1	763	23	AB171829
31	76.5	15.0	1911	22	AAO17154
32	76.5	15.0	288	23	AB171829
33	76	14.9	400	20	AB171829
34	76	14.9	1543	22	AB171829
35	75	14.7	600	23	AB171829
36	75	14.7	1024	22	AB171829
37	75	14.7	1534	22	AB171829
38	74.5	14.6	1015	22	AB171829
39	74	14.5	34	23	AB171829
40	74	14.5	449	22	AB171829
41	74	14.5	449	23	AAU74634
42	74	14.5	537	22	AB171829
43	74	14.5	837	22	AB171829
44	74	14.5	895	21	AA1548
45	74	14.5	895	22	AA1548

ALIGNMENTS

RESULT 1
AAG26460
ID AAG26460 standard; Protein; 441 AA.
XX
AC AAG26460;
XX
AC AAG26460;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30923.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

us-10-090-035-4.rag

Mon Jun 16 14:55:57 2003

PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145092.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134219.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134221.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	27-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136392.	PR	03-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	99US-0137528.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138540.	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	99US-0138847.	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	99US-0139119.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139462.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0153070.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153758.
PR	28-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0154018.
PR	29-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154039.
PR	30-JUN-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155486.
PR	06-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158222.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

Mon Jun 16 14:55:57 2003

XX	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
OS		PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD		PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
PF	06-SEP-2000.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
XX	25-FEB-2000; 2000EP-0301439.	PR	14-JUL-1999;	99US-0143824.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005.
PR	05-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144325.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	16-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	19-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148584.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150366.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

—

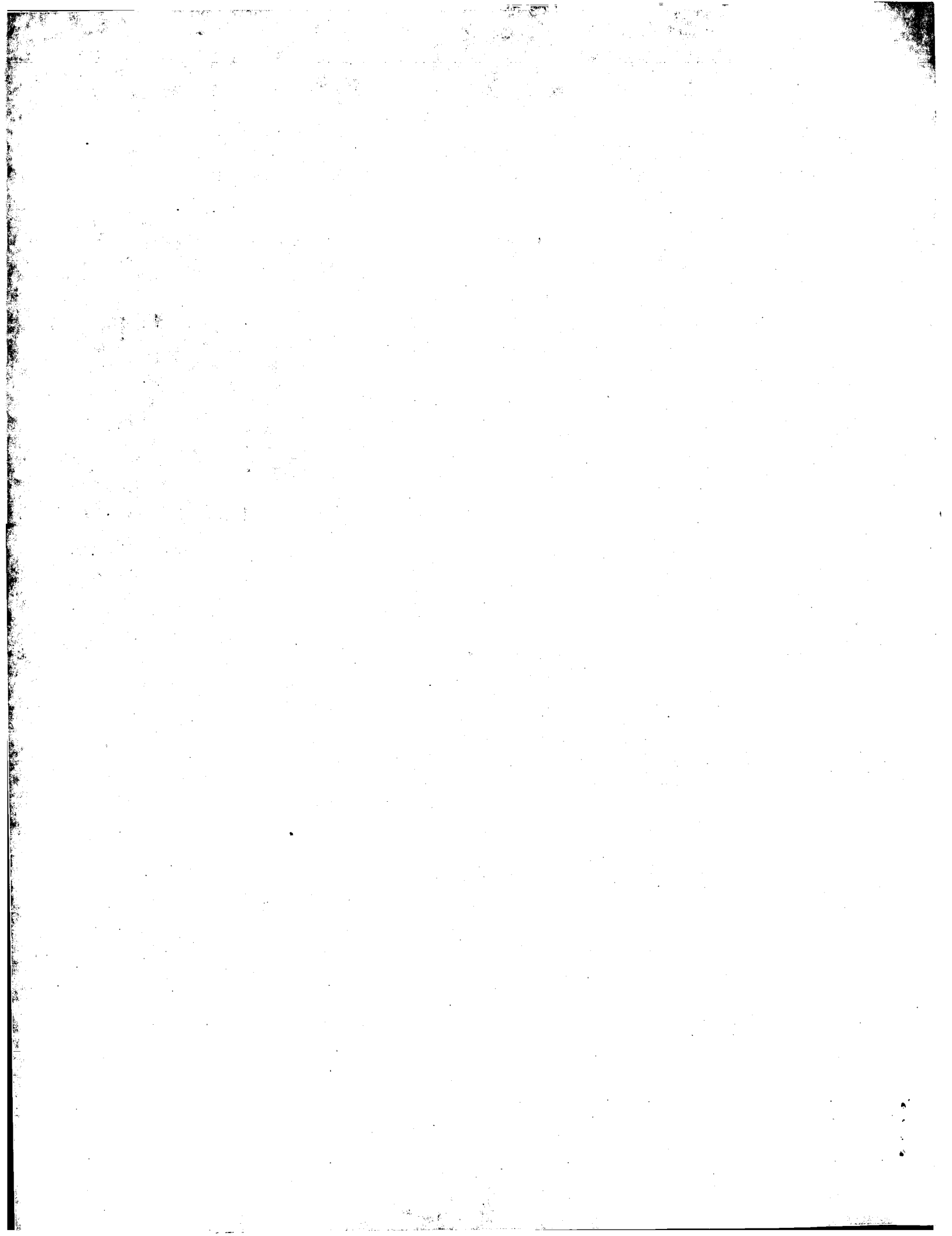
Mon Jun 16 14:55:57 2003

us-10-090-035-4.rag

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR	27-JUL-1999;	99US-01459118;
PR	27-JUL-1999;	99US-01459119;
PR	28-JUL-1999;	99US-01459521;
PR	02-AUG-1999;	99US-01463866;
PR	02-AUG-1999;	99US-01463889;
PR	02-AUG-1999;	99US-01463899;
PR	02-AUG-1999;	99US-01470389;
PR	04-AUG-1999;	99US-01472004;
PR	04-AUG-1999;	99US-01473002;
PR	05-AUG-1999;	99US-01471922;
PR	05-AUG-1999;	99US-01472609;
PR	06-AUG-1999;	99US-01473303;
PR	06-AUG-1999;	99US-01474416;
PR	09-AUG-1999;	99US-01474793;
PR	09-AUG-1999;	99US-01474935;
PR	10-AUG-1999;	99US-01481731;
PR	11-AUG-1999;	99US-01483119;
PR	12-AUG-1999;	99US-01483565;
PR	13-AUG-1999;	99US-01484851;
PR	13-AUG-1999;	99US-01486844;
PR	16-AUG-1999;	99US-01493668;
PR	17-AUG-1999;	99US-01491755;
PR	18-AUG-1999;	99US-01494266;
PR	20-AUG-1999;	99US-01497222;
PR	20-AUG-1999;	99US-01497233;
PR	20-AUG-1999;	99US-01499229;
PR	23-AUG-1999;	99US-01499902;
PR	23-AUG-1999;	99US-01499930;
PR	25-AUG-1999;	99US-01505666;
PR	26-AUG-1999;	99US-01506884;
PR	27-AUG-1999;	99US-01510665;
PR	27-AUG-1999;	99US-01510666;
PR	27-AUG-1999;	99US-01510809;
PR	31-AUG-1999;	99US-01513303;
PR	31-AUG-1999;	99US-01514348;
PR	01-SEP-1999;	99US-01519309;
PR	07-SEP-1999;	99US-01523663;
PR	10-SEP-1999;	99US-01530709;
PR	13-SEP-1999;	99US-01537558;
PR	15-SEP-1999;	99US-01540118;
PR	16-SEP-1999;	99US-01540339;
PR	20-SEP-1999;	99US-01547779;
PR	22-SEP-1999;	99US-01551339;
PR	23-SEP-1999;	99US-01553486;
PR	24-SEP-1999;	99US-01556559;
PR	28-SEP-1999;	99US-01564558;
PR	29-SEP-1999;	99US-01565696;
PR	04-OCT-1999;	99US-01571117;
PR	05-OCT-1999;	99US-01577539;
PR	06-OCT-1999;	99US-01578655;
PR	08-OCT-1999;	99US-01580229;
PR	08-OCT-1999;	99US-01582322;
PR	12-OCT-1999;	99US-01583669;
PR	13-OCT-1999;	99US-01592939;
PR	13-OCT-1999;	99US-01592994;
PR	13-OCT-1999;	99US-01592995;
PR	13-OCT-1999;	99US-01593229;
PR	14-OCT-1999;	99US-01593330;
PR	14-OCT-1999;	99US-01593331;
PR	14-OCT-1999;	99US-01596637;
PR	14-OCT-1999;	99US-01596638;
PR	18-OCT-1999;	99US-01595884;
PR	21-OCT-1999;	99US-01607411;
PR	21-OCT-1999;	99US-01607677;
PR	21-OCT-1999;	99US-01607768;
PR	21-OCT-1999;	99US-01607770;
PR	21-OCT-1999;	99US-01608115;
PR	21-OCT-1999;	99US-01608145;
PR	22-OCT-1999;	99US-01609809;
PR	22-OCT-1999;	99US-01609881;
PR	22-OCT-1999;	99US-01609889;
PR	25-OCT-1999;	99US-01614004;
PR	25-OCT-1999;	99US-01614005;

Search completed: June 6, 2003, 12:50:08
Job time : 32.5 secs




```

; OTHER INFORMATION: H37RV
US-09-103-840A-1
Query Match          9.3%; Score 53.2; DB 4; Length 4411529;
Best Local Similarity 54.6%; Pred. No. 0.028;
Matches 106; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 149 CCCGCCGGCTTCGGCCGCCACCGCGCGCGGCGGCTCCAGCAGCAGCTGCTCAAGGAGAGTT 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932728 CCCACCGGCATCGCGGCACCGCGGCACCGCGGCGGAAAGCGCGCGCGCGCGCGG 3932787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 CGAGGAGTGCACACGGTCTCACCGCGCGCGGCGGCGGCGGCGGCTCTGGTGGCGCTGGCGGCAC 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932788 CGGCGAGGCGGTAGCAGCGGTGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3932847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 CGGCGGCACCGGCTTCGTGTCGCGGAGACACGAGGTCGAAGAGGACATCAACACCTGCAC 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932848 CGGCGGACAGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3932907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 CGGCGAGGTCCACG 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3932908 CGGCGGCGCGCGG 3932921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-476-519-10
; Sequence 10, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
US-08-476-519-10
```

```

343 AGCGAGGAGAGCTTCTCGCCAGGGCT 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 ACAGCAGCAGCAGCAGCAGCGCATGGTT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          9.3%; Score 53.2; DB 4; Length 4403765;
Best Local Similarity 53.3%; Pred. No. 0.028;
Matches 112; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 147 GCCCGCGCGGCTTCGGCCGCCACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191569 GGCACCGCGCGGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1191510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 TTCGAGGAGTTCGACGGTCTCAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1191450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191509 GCGGAGCGCGCGCACACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1191450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 CACGCGCGCCACGCTTCGTGTCGCGGAGACGAGGTCGAAGAGGACATCAACACCTGC 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191449 CACGCGGGTTCGCGCGGCTGCTCCGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1191390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 ACCGCGGAGTTCACGAGCGGCGGAGGAGC 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191389 GCGGCAACGCGCGCAACGCGGCGGCAACGCG 1191360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
```


RESULT 6
US-08-476-519-1
; Sequence 1, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; City: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-476-519-1

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,519
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/281902
 FILING DATE: 28-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Bonner, Grace L.,
 REGISTRATION NUMBER: 32,963
 REFERENCE/DOCKET NUMBER: 38-21(13577)A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314)537-7286
 TELEFAX: (314)537-6047
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2334 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-476-519-1

QY	151	CGGCGCGGCTTCGGCCGCACGGGGGGCGGCTCCAGCAGCACGTCGTCAAGGAGAAAGTTCCG	210
Db	1970	CGGCGCAGCGACCAACGCGCAACGCTGATGSAGCAGTTGCGCTTCAAGCCCGACGGCG	2029
QY	211	AGGAGGTCGACACGGTCTCACGGCGCGGCGCCACCAACCACCAACCATGGTCAACCACG	270
Db	2030	CGCAGGCCGACAGCGCCCTACTTCAACGGGCGCGCACCAACCCGCCCTGGCCCTGGGGCATTG	2089
QY	271	CGGCGCCAGGGTTCTGGTGGCGGAGCACCAGGGTCGAAGAGGACATCAACACCTGCACCG	330
Db	2080	ACGGCAGCGAGTTCTGGCGACAGCGCCAGGCGCATCTACGTCGCCTTACAACGGCTGGTCCG	2149
QY	331	CGCAGGTCACCAC	341
Db	2150	CGCGGGTCGAC	2160

RESULT 7
 PCT-US95-09323-1
 ; Sequence 1, Application PC/TUS9509323
 ; GENERAL INFORMATION:
 ; APPLICANT:

RESULT 7
PCT-US95-09323-1
; Sequence 1, Application PC/TUS9509323
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09323
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-09323-1
;
; Query Match 9.0%; Score 51.8; DB 5; Length 2334;
; Best Local Similarity 54.5%; Pred. No. 0.025;
; Matches 104; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
;
; QY 151 CGGCGGCTTCGGCGCGCCAGCGGGCGGTCCAGCAGCAGCTGCTCAAGGAGAGTTGG 210
; DB 1970 CGGCCAGCAGCACCAACCGGCAAGTGATGGAGCAGTTGGCGTGGTTCAAGCGCGGCG 2029
;
; QY 211 AGGAGTTCGACAGGCTTCACGCGCGCGCCCAACCAACCAACCAACCAACCAACCAAC 270
; DB 2030 CGCAGCGCGCAGCAGCGGCTACTTCAACGCGCGCCGACCAACCAACCAACCAACCA 2089
;
; QY 271 GCGGCCACCGGCTTCGTGGTGGCGGAGCAGCAGGAGTTCGAGGAGGACATCAACACCTGC 330
; DB 2090 ACGCAGCGAGTTTCGGCGCAGCGCCAGCGCGATCTACGTGCGCTACAGCGGTGTCCG 2149
;
; QY 331 GCGAGGTCCAC 341
; DB 2150 GCGCGGTGCGAC 2160
;
; RESULT 8
; US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1
;
; Query Match 8.9%; Score 51; DB 4; Length 4411529;
; Best Local Similarity 52.6%; Pred. No. 0.078;
; Matches 111; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
;
; QY 132 GAGGTGAGTTCGTGGCGCGCGCGCTTCGCGCCGCCACAGCGGCGGTCCAGCAGCAC 191

;
; 3749067 GAGTCGGGCACTGTAATCGCTGGAACATCGTCAACTGGGCAACCTCGGCAGCTACAC 3749008
;
; QY 192 GTGCTCAAGGAGAGTTTCGAGGAGGTCGACACAGCTCTCAGCGCGCGCGCCAAACCAAC 251
; DB 3749007 CTCGGCTTCGCCAACGTCGCGGACGTCACCTGGCGGGCAACCTCGGCAACCTCAAC 3748948
;
; QY 252 CACCACCATGTCACACACGCGCGCGGCTTGTGTGTGCGGAGACCAACGAGGTCGAAGAG 311
; DB 3748947 CTCGGCGGTGCAACCTTCGGCGGGCAAGACCTTGGGCTTGGGCAACCTTCGGGGAGCGCAAC 3748888
;
; QY 312 GACATCAACACCTTCACCGCGGAGGTCCACG 342
; DB 3748887 GTCGGGTTTCGGCAACCTTCGGCCACCGCAATG 3748857
;
; RESULT 9
; US-08-440-856A-9
; Sequence 9, Application US/08440856A
; Patent No. 5750873
; GENERAL INFORMATION:
; APPLICANT: DELLAPORTA, STEPHEN L.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
; TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE. N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,856A
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A.
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 05463-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1517
; TELEFAX: (202) 887-0763
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-440-856A-9
;
; Query Match 8.7%; Score 50.2; DB 1; Length 1288;
; Best Local Similarity 43.9%; Pred. No. 0.05;
; Matches 112; Conservative 20; Mismatches 123; Indels 0; Gaps 0;
;
; QY 116 GGACTACTGCTCGGAGAGTTCGAGGAGGTCGAGGAGGTCGAGGAGGTCGAGGAGGTCGAGG 175
; DB 754 GAACGCGCTTCGGAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTCG 813
;
; QY 176 CGGCTTCACGACGTCGTCGAGGAGGTCGAGGAGGTCGAGGAGGTCGAGGAGGTCGAGGAGG 235
; DB 814 CGTGCACCCGCCATGCTCATCAACGCTGCGCGGAGGTCGAGGAGGTCGAGGAGGTCGAGGAG 873
;
; QY 236 CGGCGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 295
; DB 874 CGMCGCCGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 933

Mon Jun 16 14:55:57 2003

```

NAME/KEY: polva_site
LOCATION: 7314..7319
FEATURE:
NAME/KEY: prim_transcript
LOCATION: 2072
FEATURE:
NAME/KEY: promoter
LOCATION: 1..2155
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 2042..2049
FEATURE:
NAME/KEY: mRNA
LOCATION: 2153..7079
US-08-250-848-2

Query Match      8.6%; Score 49.2; DB 2; Length 7559;
Best Local Similarity 49.6%; Pred. No. 0.094; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 110 GGAGGTGACTACTGTCGAGAGAGTGTGAGTGTGCGTGGCCCGCGGCTTCGGCGCCCA 169
Db 2490 GAAGCTGGGGAGCTGGCGCCCAAGCTCAGGGGCTGGCCCGCGGAGGCTCCTCGT 2549

QY 170 CGCGCGGGGTCACACACAGTGTCAAGAGAGTTCGAGGAGGTTCGACACGTCCTC 229
Db 2550 GCGAGAGTCCATCTCGCATGCTCAACCTGCGCAACCTGCGCAAGGTGGTTCGCGGAGGAGTGCATCGC 2609

QY 230 ACGCGCGGGCGCCACACACACCATGCTCACCACGCGGCGCCAGGCTTCGTGT 289
Db 2610 GCACCGCGCGCCAAACAGCAAGCTCAAGAAAGTGGTTCGCGGAGGAGGCTCGGCCAC 2669

QY 290 GCGCGAGACAGGTCGAGAGACATCAACACCTGACCGGCGGAGGTCACGAGGCGAG 349
Db 2670 CACCGAGTCCGATCGAGAGAGCTCAAGCGCTCGTGTGCGAGGTCGGCAAGTCCCC 2729

QY 350 GGAGAGTTCCTCG 363
Db 2730 CGAGAGGTTCTCG 2743

RESULT 11
US-08-440-856A-1
; Sequence 1, Application US/08440856A
; Patent No. 5750873
; GENERAL INFORMATION:
; APPLICANT: DELLAPORTA, STEPHEN L.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
; TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; CITY: 2000 PENNSYLVANIA AVE. N.W.
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,856A
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A.
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 05463-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1517
; TELEFAX: (202) 887-0763

TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-440-856A-1

Query Match      8.5%; Score 48.6; DB 1; Length 1236;
Best Local Similarity 49.4%; Pred. No. 0.1; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 116 GGACTACTGTCGAGGAGGTGTGAGTGTGCGTGGCCCGCGGCTTCGGCGCCCA 175
Db 743 GAACCGCGCTTCGAGGTGCGCGCGCACGCGGTCAACTGCTCTCGGCTTCGG 802

QY 176 CGCGGTCCAGCAGCACCTGCTCAAGAGAGTTCGAGGAGGTTCGACACGCTCTCACGCGC 235
Db 803 CGTCCGACGCCCATGCTCATCAACGCTGCGCGCCAGGCGCCACGACGCCACGCCGA 862

QY 236 CGCGCGCAACACACACACCATGCTGTCACCGCGGCGCCAGGCTTCGTGTTCGCGCA 295
Db 863 CGCGCGCAGGAGCTCGACCTCGACCTGCGAGTCCACGTCGCCAGCGCAGGAGGTGA 922

QY 296 GACCAGGTCGAAGAGGAGATCAACACCTGACCGCGGAGGTTCACGAGCGCAGGAGAG 355
Db 923 GAAGATGAGGAGGTGTGAGGCGCTGCGCCACGCTCAAGGGGCCCCACGCTCAGGCGCCAG 982

QY 356 CTTCTCGCCAGGC 370
Db 983 GGACATCGCGAGGC 997

RESULT 12
US-08-133-711-44
; Sequence 44, Application US/08133711
; Patent No. 5525463
; GENERAL INFORMATION:
; APPLICANT: Zolig, Werner
; TITLE OF INVENTION: Methods and reagents for detection of
; TITLE OF INVENTION: pathogens using superoxide dismutase gene
; TITLE OF INVENTION: targeting
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: U.S.A.
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,711
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92810780.4
; FILING DATE: 13-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Patricia S. Rocha
; REGISTRATION NUMBER: 31,054
; REFERENCE/DOCKET NUMBER: 4095/95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235 5000
; TELEFAX: (201) 235 3500
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Actinomyces viscosus/SOD gene
US-08-133-711-44

Query Match 8.4%; Score 48; DB 1; Length 491;
Best Local Similarity 50.9%; Pred. No. 0.13;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 147 GCCCGCGCGGCTTCGCGCGCCAGCGCGCGCGCTCCAGCAGCAGTGTCTCAAGAGAAG 206
DB 51 GCGCTGGAGGCTTCGCGCGCGCGCGCGAGGAGCGGACCTGGGTGCATCAACCTGTGG 110
QY 207 TTCGAGGAGTTCGACACGCTTCACGCGCGCGCGCGCCAAACCCACCCACCATGCTCAC 266
DB 111 GAGAGAGCTTCGCTTCAACCTGGCGCGCGCCACACCACTCGTGTCTTGGAGAAG 170
QY 267 CACGCGCGCGCGCTTCGCTGTGCGCGAGACCGAGCTTCCTTCGCCAGGCG 370
DB 171 CTCTCCCGCAACGCGCGCGCGCGAGCGCGGAGCTGCGGAGGCCATCAAGGACTCC 230
QY 327 ACCGCGGAGTTCACGAGCGCGAGGAGAGCTTCCTTCGCCAGGCG 370
DB 231 TTCGCTCTTTCGAGAAGTTTCAGGCGCGAGTTTCACCGCGCACCGC 274

RESULT 13

US-08-440-856A-2

Sequence 2, Application US/08440856A

Patent No. 5750873

GENERAL INFORMATION:

APPLICANT: DELLAPORTA, STEPHEN L.

TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING

TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 PENNSYLVANIA AVE. N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,856A

FILING DATE: 15-MAY-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MILLMAN, ROBERT A.

REGISTRATION NUMBER: 36,217

REFERENCE/DOCKET NUMBER: 05463-20001.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1517

TELEFAX: (202) 887-0763

TELEX: 706141

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1187 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-440-856A-2

Query Match

Best Local Similarity 47.6%; Score 48; DB 1; Length 1187;

Matches 141; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 79 ACCAAGCGTCGCGCACCAATGCTTACTACCAAGAGGTGGACTACTGCTCGGAGAGGTGA 138
DB 535 AGCAGCGCGCGCTCGCCATGACCCAGCGCGCGCGCAGCATCATCTCGGTGCGCCAGCG 594
QY 139 GGTGCGTGGCCCCCGCGGCTTCGCGCGCCACGCGCGCGCTCCAGCAGCAGCAGCTCTCA 198
DB 595 TCGCGCGCGTCTCGCGCGCGCTCGCGCGCGCTACACCGCTCCCAAGCAGCGCCATCG 654
QY 199 AGGAGAAGTTCGAGGAGGTTCGACACGCTCTCACCGCGCGCGCGCGCCACCAACACAC 258
DB 555 TGGGGCTACCAAGAACCGCGCTTCGAGCTCGCGCGCGCGCGCGCATCCGGTCAACTGCA 714
QY 259 ATGCTACCAACGCGCGCGCGCTTCGCTGTGCGCGAGACGAGGTGGAAGAGGACATCA 318
DB 715 TCTCCCTTCGCGCGTCCGCGCGCGCTTCATCAACGCTTCGCGCGCGCGCGCGCGCG 774
QY 319 ACACCTGACCGCGCGGAGTTCACGAGCGCGAGGAGAGCTTCCTCGCGCGCGCGCGTAA 374
DB 775 CCTCCACCG 830

RESULT 14

US-09-103-840A-2

Sequence 2, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 8.3%; Score 47.8; DB 4; Length 4403765;

Best Local Similarity 49.8%; Pred. No. 0.34;

Matches 121; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 151 CGCGCGGCTTCG 210
DB 3921743 CG 3921802
QY 211 AGGAGTTCGACACGCTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921802
DB 3921803 GCGGGATCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921862
QY 271 GCGGCGCGCGCTTCGCTGTGCGCGAGACGAGGTGGAAGAGGACATCAACACCTGCG 330
DB 3921863 CCGGCAACACAGGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3921922
QY 331 GCGAGTTCACGAGCG 390
DB 3921923 GAGCGGTGAACG 3921982
QY 391 CGG 393

Query Match

Best Local Similarity 47.6%; Score 48; DB 1; Length 1187;

Matches 141; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

```

RESULT 15
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent NO. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

```

Query Match	8.28;	Score 47.2;	DB 4;	Length 1925;
Best Local Similarity	49.6%;	Pred. NO. 0.21;		
Matches 121;	Conservative 0;	Mismatches 123;	Indels 0;	Gaps 0;
QY	109	AGGAGGTGGACTACTCTCGGAGGAGGTGAGGTGCGCTGGCCCGCGCGCTTCGCGCCGCC	168	
Db	924	ACGGGAGGACGAGGACGGGGAGGACGGGGAGGACGAGGACGGGGAGGACGGGGAGGACG	865	
QY	169	ACGGGGGGGGTCCACGACACGTGTCGAGGAGAGTTTCGAGGAGGTTCGAGGAGTTCGACACGTCT	228	
Db	864	AGGACGGGGAGGACGGGGAGGAGGACGAGGACGGGAGGACGGGGAGGAGGACGAGGACG	805	
QY	229	CACGCGCGCGGCCCAACCCACCCACCATGTGTACACGCGCGGCCACGGCTTCGTGG	288	
Db	804	GGGAGGACGGGGAGGACGAGGACGGGGAGGACGGGGAGGACGAGGACGGGGAGGACGGGG	745	
QY	289	TCCGGGACACAGGTCGAGAGGACATCAACACTTCACCGCGGAGGTTCACGAGCGCA	348	
Db	744	AGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGG	685	
QY	349	GGGA	352	
Db	684	AGGA	681	

Search completed: June 16, 2003, 03:58:13
Job time : 81.4096 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 23:41:35 ; Search time 188.272 Seconds
(without alignments)
6865.849 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.6	10.9	390	13	AAQ21833
2	62.6	10.9	390	14	AAQ36859
3	62.6	10.9	390	22	AAQ76910
4	62.6	10.9	390	24	AAQ72775
5	53.2	9.3	4403765	22	AAI99683
6	53.2	9.3	4411529	22	AAI99682
7	52.8	9.2	1128	20	AAQ10560
8	52.8	9.2	1128	21	AAQ10560
9	52.8	9.2	1128	21	AAQ46026

c	10	52.8	9.2	1128	24	AAQ27892	Human G-protein co
c	11	52.8	9.2	1128	24	AAQ27892	Human DNA for pote
c	12	52.8	9.2	109519	22	AAQ8049	Micromonospora DNA
c	13	52.8	9.2	114955	20	AAQ3491	Human adenosine A1
c	14	51.8	9.0	2244	17	AAQ10429	Mature isomylase
c	15	51.8	9.0	2244	19	AAQ23640	Flavobacterium iso
c	16	51.8	9.0	2334	17	AAQ10428	isoamylase gene.
c	17	51.8	9.0	2334	19	AAQ23639	Full length Flavob
c	18	51.4	9.0	3849	22	AAQ25795	S. chrysomallus ac
c	19	51	8.9	4411529	22	AAI99682	Mycobacterium tube
c	20	50.8	8.9	3059	8	AAQ70566	Sequence encoding
c	21	49.8	8.7	522	21	AAQ24813	Wheat inositol 1.3
c	22	49.6	8.6	685	24	ABQ78045	Wheat SCIP-1 ortho
c	23	49.4	8.6	65140	22	AAQ17184	Streptomyces nous
c	24	49.4	8.6	125401	22	AAQ17186	Streptomyces nous
c	25	49.2	8.6	1171	22	AAQ44082	Oryza sativa perox
c	26	49.2	8.6	6781	19	AAQ45824	Maize phosphoenolp
c	27	49.2	8.6	7559	20	AAQ82011	Maize Ts2 CDNA nuc
c	28	49	8.5	1470	23	AAQ54218	Streptomyces nous
c	29	48.6	8.5	1236	16	AAQ45060	Maize phosphoenolp
c	30	48.2	8.4	27541	22	AAQ17185	Maize phosphoenolp
c	31	48.2	8.4	125401	22	AAQ17186	Streptomyces nous
c	32	48	8.4	490	15	AAQ62127	Streptomyces nous
c	33	48	8.4	1725	21	AAQ10503	Superoxide-dismuta
c	34	48	8.4	2218	21	AAQ10516	Trehalose-releasin
c	35	47.8	8.3	4403765	22	AAI99683	Mycobacterium tube
c	36	47.4	8.3	14805	22	ABQ3809	Streptomyces galli
c	37	47.2	8.2	1464	24	ABQ90041	M. capsulatus gene
c	38	47	8.2	1260	11	AAQ05666	TGF-57-pseudomonas
c	39	47	8.2	1260	11	AAQ06127	Sequence encoding
c	40	47	8.2	1260	18	AAQ72116	TGF-57-PE40 gene d
c	41	47	8.2	1260	17	AAQ76879	Wild type TGF-alpha
c	42	47	8.2	2291	20	AAQ16859	Nitric oxide synth
c	43	47	8.2	2668	21	AAQ00337	Wheat raffinose sy
c	44	47	8.2	3616	17	AAQ16858	Nitric oxide synth
c	45	47	8.2	4089	17	AAQ16857	Bovine endothelial

ALIGNMENTS

RESULT 1
AAQ21833
ID AAQ21833 standard; DNA; 390 BP.
AC AAQ21833;
XX
DT 08-JUN-1992 (first entry)
XX
DE Randomising oligonucleotide used in SPERT mRNA prepn.
DE
KW Systematic polypeptide evolution by reverse translation; SPERT;
KW ligand binding; ss.
XX
OS Synthetic.
XX
PN WO9202536-A.
XX
PD 20-FEB-1992.
XX
PF 01-AUG-1991; 91WO-US05463.
XX
PR 02-AUG-1990; 90US-0561968;
XX
XX (COLS) UNIV OF COLORADO.
XX
XX Gold L, Tuerk C;
XX
XX WPI; 1992-080018/10.
XX
XX New method of systematic polypeptide evolution by reverse
XX translation - by linking each polypeptide in sample mixt. to
XX individualised mRNA allowing further synthesis of selected

PT	polypeptide(s)	
XX	Example; Page 55; 102pp; English.	
PS		
CC	The sequence is that of an example randomising oligonucleotide which	
CC	is used in the prepn. of mRNA encoding candidate polypeptides for the	
CC	method of systematic polypeptide evolution by reverse translation	
CC	(SPERT). The method provides a rapid way of isolating and identifying	
CC	polypeptide ligands which bind to target mols. The polypeptide ligands	
CC	can be used in e.g. assay methods, diagnostic procedures, cell sorting,	
CC	as activators or inhibitors of target mol. function, as probes, as	
CC	sequestering agents, drug delivery vehicles, modifiers of hormone	
CC	action and as catalysts. See also AAQ21840-021832.	
XX		
SQ	Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;	
	Query Match 10.9%; Score 62.6; DB 13; Length 390;	
	Best Local Similarity 52.0%; Pred. No. 0.0016;	
	Matches 140; Conservative 0; Mismatches 129; Indels 0; Gaps 0;	
QY	103 ACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCCCGCGGCTTCG 162	
Db	109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168	
QY	163 GCCGCCACGCGCGGCTCCACGACGACGACGACGACGACGACGACGACGACGACG 222	
Db	169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228	
QY	223 CGGTCTCACGCGCGCGCCCAACGACGACGACGACGACGACGACGACGACGACG 282	
Db	229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288	
QY	283 TCGTGGTCCGCGAGACGACGAGGTGCGAGAGGACATCAACACCTCCACCGCGGAGTCCAG 342	
Db	289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348	
QY	343 AGCCGAGGAGAGCTTCTCGCCAGGGCT 371	
Db	349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377	
RESULT 2		
AAQ36859		
ID	AAQ36859 standard; DNA; 390 BP.	
XX		
AC	AAQ36859;	
XX		
DT	22-JUN-1993 (first entry)	
XX		
DE	PCR primer for 5' fixed sequence contg. T7 promoter and RBS.	
XX		
KW	Systematic peptide evolution by reverse translation; SPERT; ligand;	
XX	specific; inhibitors; probes; assay; cell sorting; ss.	
OS	Synthetic.	
XX		
PN	WO9303172-A.	
XX		
PD	18-FEB-1993.	
XX		
PF	31-JAN-1992; 92WO-US00801.	
XX		
PR	01-AUG-1991; 91US-0739055.	
XX		
XX	(UYRE-) UNIV RES CORP.	
XX		
PI	Gold L, Pribnow D, Smith JD, Tuerk C;	
XX		
DR	WPI; 1993-076529/09.	
XX		
PT	Systematic polypeptide evolution by reverse translation - used	
PT	for prodn. of polypeptide ligand specific for desired target	
PT	molecule	

XX	Example 1; Page 84; 98pp; English.	
PS		
CC	SPERT is used to select novel polypeptides that bind the antibody	
CC	of the epitope commonly recognised by the antisera from autoimmune	
CC	mice which are the fl progeny of a cross of NZB and NZW parents	
CC	(Portanova et al., J. Immunol. 144, 4633, 1990). The known epitope	
CC	consists of ca. 10 amino acids at the N-terminus of the histone H2B	
CC	protein. To make mRNA encoding candidate polypeptides a 5' fixed	
CC	sequence composed of a T7 promoter sequence and a ribosome binding	
CC	site which is recognised by both prokaryotic and eukaryotic ribosomes,	
CC	terminating in a restriction endonuclease site is synthesised and cloned	
CC	using a number of oligonucleotides (example shown). A 3' fixed sequence	
CC	is placed into a restriction site to provide an mRNA encoding the C-	
CC	terminal trailer sequence of ca. 100 nucleotides lacking stop codons.	
CC	In addition, a 3' primer annealing site is provided so that cDNA	
CC	synthesis can be accomplished on the mRNA recovered from partitioned	
CC	ribosome complexes. See also AAQ36845-63.	
XX		
SQ	Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;	
	Query Match 10.9%; Score 62.6; DB 14; Length 390;	
	Best Local Similarity 52.0%; Pred. No. 0.0016;	
	Matches 140; Conservative 0; Mismatches 129; Indels 0; Gaps 0;	
QY	103 ACTACACGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGGTGGCCCGCGGCTTCG 162	
Db	109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168	
QY	163 GCCGCCACGCGCGGCGCTCCACGACGACGACGACGACGACGACGACGACGACGACG 222	
Db	169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228	
QY	223 CGGTCTCACGCGCGCGCCCAACGACGACGACGACGACGACGACGACGACGACG 282	
Db	229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288	
QY	283 TCGTGGTCCGCGAGACGACGAGGTGCGAGAGGACATCAACACCTCCACCGCGGAGTCCAG 342	
Db	289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348	
QY	343 AGCCGAGGAGAGCTTCTCGCCAGGGCT 371	
Db	349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377	
RESULT 3.		
AAQ76910		
ID	AAQ76910 standard; DNA; 390 BP.	
XX		
AC	AAQ76910;	
XX		
DT	29-MAY-2001 (first entry)	
XX		
DE	Sequence containing a 120 repeat of ACG flanked by fixed fragments.	
XX		
KW	Ligand isolation; systemic polypeptide evolution by reverse translation;	
XX	SPERT; ss.	
OS	Synthetic.	
XX		
PN	US6194550-B1.	
XX		
PD	27-FEB-2001.	
XX		
PF	23-NOV-1998; 98US-0197649.	
XX		
PR	31-JAN-1992; 92US-0829461.	
XX		
PR	02-AUG-1990; 90US-0561968.	
XX		
PR	01-AUG-1991; 91US-0739055.	
XX		
PA	(GOLD/) GOLD L.	
PA	(TUERK/) TUERK C.	


```

Query Match          9.3%; Score 53.2; DB 22; Length 4403765;
Best Local Similarity 53.3%; Pred. No. 0.28;
Matches 112; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 147 GCCCGCCGCGCTTCGCGCCGCGCCAGCGCGCGGCTCCAGCAGCAGCAGCTCGTCAAGGAGAAG 206
Db 1191569 GGCACCGCGCGCTTCGCGCCGCGCCAGCGCGCGGCTCAGCGCGCGGCGCGGCTCCAG 1191510

QY 207 TTCAGAGAGTTCGACAGCGTCTCACGCGCGCGCGCCCAACACCAACCAACCAACCAACCAAC 266
Db 1191509 GCGGAGCGCGCGCGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGC 1191450

QY 267 CACGCGCGCCACGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 326
Db 1191449 GACGCGCGGTGCGCGCGCTGCTGCGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191390

QY 327 ACCGCGGAGGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 356
Db 1191389 GCGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191360

RESULT 6
AAI99682
ID AAI99682 standard; DNA; 4411529 BP.
XX
AC AAI99682;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ
XX
PS Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
```

```

103 ACTACGAGGAGTGGACTACTCTCGGAGGAGGTGAGTGGTGGCGCGCGCGCGCTTCG 162
Db 109 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 168

QY 163 GCCCGCCGCGCGCTTCGCGCCGCGCCAGCGCGCGGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAGTTCGACA 222
Db 169 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 228

QY 223 CGGTCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 282
Db 229 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 288

QY 283 TCGTGGTGGCGGAGACGAGGTTCGAGAGGAGCATCAACACCTGCACCGCGCGAGGTTCGACG 342
Db 289 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 348

QY 343 ACGCGAGGAGGAGTTCCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 371
Db 349 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 377

RESULT 5
AAI99683/C
ID AAI99683 standard; DNA; 4403765 BP.
XX
AC AAI99683;
XX
DT 15-JAN-2002 (first entry)
XX
DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US6294328-B1.
XX
PD 25-SEP-2001.
XX
PF 24-JUN-1998; 98US-0103840.
XX
PR 24-JUN-1998; 98US-0103840.
XX
PA (GENO-) INST GENOMIC RES.
XX
PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI; 2001-647261/74.
XX
PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ
XX
PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
```



```
QY 337 TCCACGAGCGCAGGAGACG 356
Db 482 TCCAGCACTGGCGGGAAGCG 463

RESULT 10
AAD27892/C
ID AAD27892 standard; DNA; 1128 BP.
AC AAD27892;
XX
XX
DT 31-MAY-2002 (first entry)
XX
DE Human G-protein coupled receptor 14266 DNA.
XX
KW Human; G-protein coupled receptor; GPCR 14266; haematopoietic disorder;
KW neutrophil deficiency disorder; splenomegaly; pulmonary hypertension;
KW colon; diarrhoea; hepatic injury; idiopathic inflammatory bowel disease;
KW uterus; endometriosis; brain; acute meningitis; multiple sclerosis;
KW T-cell; systemic lupus erythematosus; skin; vitiligo; heart failure;
KW angina pectoris; atherosclerosis; haemolytic anaemia; thymic cyst; ds.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1128
FT /tag= a
FT /product= "GPCR 14266"
FT sig_peptide 1..102
FT /tag= b
FT mat_peptide 103..1125
FT /tag= c
FT /product= "Mature GPCR 14266"
XX
XX
PN WO200212344-A2.
XX
XX
PD 14-FEB-2002.
XX
XX
PF 08-AUG-2001; 2001WO-US24835.
XX
PR 09-AUG-2000; 2000US-0634392.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Welch NS;
XX
XX
DR WPI; 2002-217181/27.
XX
P-PSDB; AAE14597.
XX
XX
PT Identifying compound that binds to and/or modulate activity of the
PT polypeptide for treating splenomegaly, emphysema and multiple
PT sclerosis, comprises using human G-protein coupled receptor 14266
PT polypeptide -
XX
XX
Claim 1; Fig 2; 115pp; English.
XX
XX
XX
The invention relates to a method of identifying an agent that binds to
and/or modulates activity of human G-protein coupled receptor (GPCR)
14266. The invention also provides a method for modulating the level
or activity of the GPCR in cells in vivo or in vitro, or in a subject
predisposed to having a haematopoietic or neutrophil deficiency
disorders. Detecting 14266 polypeptides or nucleic acids in specific
biological samples is useful in disease diagnosis. Modulating level or
activity of the GPCR polypeptide or polynucleotide is useful for treating
diseases related to 14266 receptor malfunction e.g., disorders involving
spleen (e.g. splenomegaly, neoplasms); lung (e.g. pulmonary hypertension,
emphysema); colon (e.g. diarrhoea and dysentery, idiopathic inflammatory
bowel disease); liver (e.g. hepatic injury, viral hepatitis); uterus and
endometrium (e.g. endometriosis, endometrial polyps); brain (e.g.
acute meningitis, multiple sclerosis, spinocerebellar degeneration);
T-cells (e.g. systemic lupus erythematosus, polyarthritis nodosa); skin
(e.g. vitiligo, acanthosis nigricans); heart (e.g. heart failure,
```

```
CC angina pectoris); blood vessels (e.g. atherosclerosis, Raynaud disease,
CC thalassaemia syndromes, haemolytic anaemia) and disorders of the thymus
CC (e.g. thymic cysts, Hodgkin's disease). The present sequence is human
CC GPCR 14266 DNA.
XX
```

SQ Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;

Query Match 9.2%; Score 52.8; DB 24; Length 1128;
Best Local Similarity 52.7%; Pred. No. 0.15;
Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

```
QY 97 TGGCTTACTACGAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTGCGTGGCCGCGG 156
Db 721 TGGCGGCGCGCTGGCGGTGGCGCGCGCGCGCGTGAAGGTCCAGTCTGCTGACGCGG 662
QY 157 GCTTTCGGCGCCACGCGCGCGGTCCAGCAGCAGCGTCAAGGAGAGTTCGAGGAGG 216
Db 661 GCACGAGCGCGCGCGCGCGCATTTGGCGGCGGTGGATGAAGAGCAGCGCGGAGG 603
QY 217 TCACAGCGGTCTCACGCGCGCGCGCGCAACACACACACCATGTTGTCACCGCGCGG 276
Db 602 TAGACGAGGTGCGTGGCGCGCGCGCACACACGCGCAGCAGGAGGAGCCGCGCGG 543
QY 277 ACGGCTTCGTGGTGGCGGAGCAGCGGTCCAGAGGAGCATCAACACTGCACCGCGGAG 336
Db 542 GGGCGCGCTGGCGCGCGCTCTCCAGGCGCGACGCGCGTCTCTGCTGCGCACCGCGG 483
QY 337 TCCACGAGCGCAGGAGAGC 356
Db 482 TCCAGCACTGGCGGGAAGCG 463
```

RESULT 11

AAS98049/c
ID AAS98049 standard; DNA; 1128 BP.

AC AAS98049;

XX 12-MAR-2002 (first entry)

DE Human DNA for potential G protein-coupled receptor #7.

XX Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
KW depression; epilepsy; macular degeneration; lymphoma; melanoma;
KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KW psoriasis; rheumatoid arthritis; schizoprenia; ulcerative colitis;
KW tuberculosis; cognition disorder; memory disorder; anorexia;
KW hormonal release disorder; cardiovascular activity disorder;
KW pain perception disorder; obesity; diabetes; obesity;
KW diabetes; hyperlipidaemia; stroke; gene therapy.

XX Homo sapiens.

XX WO200185791-A1.

XX 15-NOV-2001.

XX 11-MAY-2001; 2001WO-US15332.

XX 11-MAY-2000; 2000US-203217P.

XX 18-MAY-2000; 2000US-205945P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;

XX WPI; 2002-066595/09.

XX Novel G protein-coupled receptor polypeptides including galanin
PT receptor polypeptides useful for identifying modulators that are useful

PT for treating Alzheimer's disease, psoriasis, melanoma, multiple
PT sclerosis, stroke
XX
PS
XX
XX
CC The invention relates to an isolated polypeptide encoded by a
CC nucleic acid molecule that is at least 80% identical to the G
CC protein-coupled (GPCR) polynucleotides included in the specification.
CC Also included are probes based on the GPCR sequences (including
CC antisense probes), a host cell comprising an expression vector comprising
CC the GPCR sequence, antibodies raised against the polypeptides,
CC and methods of identifying modulators of the polypeptides. The
CC polypeptides are useful for identifying modulator compounds which
CC function as modulators, activators, repressors, agonists or antagonists
CC of the novel GPCR polypeptides including the G4U polypeptide. The
CC antibodies and nucleic acid probes as described above can be used to
CC detect the presence of the polypeptides and nucleic acids and are used to
CC diagnose a variety of diseases or disorders in which GPCRs are involved
CC e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,
CC atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,
CC chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,
CC depression, epilepsy, macular degeneration, lymphoma, melanoma,
CC multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,
CC psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,
CC tuberculosis and many other diseases listed in the specification. The
CC probes and antibodies are also useful for diagnosing cognition and memory
CC disorders, anorexia, hormonal release disorders, cardiovascular activity
CC disorders, pain perception disorders, obesity, diabetes, Alzheimer's
CC disease. Preferably, compounds that decrease or increase
CC the expression of galanin receptor (GAL4) can be used to treat obesity,
CC diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is
CC useful for treating the above mentioned disorders by gene therapy
CC techniques. The present sequence is a novel GPCR polynucleotide of the
CC invention.
XX
SQ Sequence 1128 BP; 116 A; 437 C; 395 G; 180 T; 0 other;
Query Match 9.2%; Score 52.8; DB 24; Length 1128;
Best Local Similarity 52.7%; Pred. No. 0.15;
Matches 137; Conservative 0; Mismatches 122; Indels 1; Gaps 1;
QY 97 TGCGTTACTACAGGAGGTGGACTACTGTCGAGGAGGTGAGTGGCGCCCGCGCG 156
Db 721 TGGCGCGCGCTGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 562
QY 157 GCTTCG 216
Db 661 GCACGAGGCG 603
QY 217 TCGACACGCGTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
Db 602 TAGACGAGGTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 543
QY 277 ACAGCTTCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 336
Db 542 GGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
QY 337 TCCACGAGCG 356
Db 482 TCCAGCACTGGCGGGAAGGC 463
RESULT 12
AAS08693
ID AAS08693 standard; DNA; 109519 BP.
XX AAS08693;
AC AAS08693;
XX
XX 26-SEP-2001 (first entry)
DT
XX Microspora DNA encoding biosynthetic enzymes for Everninomycin.
DE
DE Everninomycin; antibiotic; bottle-neck gene; orthomycin;
KW

fermentation; ds.
KW XX OS Micromonospora carbonacea var. africana.
XX
XX
FH Key Location/Qualifiers
FT CDS complement (132..1382)
FT /tag= a
FT /product= "EvdA"
FT complement (1389..1394)
FT /tag= b
FT complement (1490..2611)
FT /tag= c
FT /product= "EvdB"
FT complement (2618..2622)
FT /tag= d
FT complement (2622..3860)
FT /tag= e
FT /product= "EvdC"
FT complement (3867..3870)
FT /tag= f
FT 4143..5312
FT /tag= g
FT /product= "EvdD"
FT 4134..4138
FT /tag= h
FT 5309..6235
FT /tag= i
FT /product= "EvdE"
FT 6232..7275
FT /tag= j
FT /product= "EvdF"
FT 6226..6229
FT /tag= k
FT 7272..8327
FT /tag= l
FT /product= "EvdG"
FT 8342..9364
FT /tag= m
FT /product= "EvdH"
FT 8333..8336
FT /tag= n
FT complement (9463..10224)
FT /product= "EvdI"
FT complement (10232..10235)
FT /tag= o
FT 10424..11176
FT /tag= q
FT /product= "EvdJ"
FT 12027..12455
FT /tag= r
FT /product= "EvdK"
FT partial
FT /note= "No start codon"
FT complement (12108..13022)
FT /tag= s
FT /product= "EvdL"
FT complement (13027..13030)
FT /tag= t
FT complement (14410..15363)
FT /tag= u
FT /product= "EvdR"
FT complement (15369..15373)
FT /tag= v
FT complement (15380..16414)
FT /tag= w
FT /product= "EvdS"
FT complement 16419..17873
FT /tag= x
FT /product= "EvdT"
FT complement (17870..18934)
FT /tag= y
FT /product= "EvdU"

FT	CDS	19374..20906	/**tag= z	FT	CDS	106	ACCAGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGTGGCCCGCGCGGCTTCGGCC	165
FT	FT	/product= "EvrE"	complement (43807..43811)	FT	RBS	11812	AGCGGCGCGGTGACGGCGACCGGCGCGGTGCGGCTGCTGAACTGCGGCTCGGCTCGGCGC	11871
FT	CDS	21064..22542	/**tag= aa	FT	FT	166	GCCACGCGCGCGGCTGCCAGACGACGTCGTCAGGAGAGTTCGAGGAGGTGCGACACGG	225
FT	FT	/product= "EvrF"	complement (43799..44866)	FT	FT	11872	TGCGCACCTCCCTCGCGGAGTTGCTCCAGACAGTGTCCAGGTGGCGGCTGTCGAGCTGG	11931
FT	RBS	21056..22542	/**tag= ab	FT	FT	226	TCTACGCGCGCGGCGCCAAACACACACCATGGTCCACGCGCGCGGCGGCTTCG	285
FT	FT	22748..24172	/**tag= ac	FT	FT	11932	AGGTATATCGGGAGGTCCGCGGAGTTCGACCATCGGGGCAACTGGCTCGACACACCC	11991
FT	CDS	/**tag= ac	complement (45014..45760)	FT	FT	286	TGGTGGCGGAGACCGAGGTTCGAGAGGAGATCAACACCTGCACCGCGGAGGTCCAGGAGC	345
FT	FT	/product= "EvrG"	complement (45767..45770)	FT	FT	11992	TGCGCGGAGACCGCTGGGCTGGCAGGCGGATCAGCTCCCGGCGGCGGCTCGGCGCAGT	12051
FT	FT	22736..22740	/**tag= ad	FT	FT	346	GCAGGAGAGCTTCTCTCGCAGGCTTAAGTACGCGCGCGGCGGCGGCTCCAGCGCGC	405
FT	CDS	/**tag= ae	complement (24177..25223)	FT	FT	12052	GCTGGAGGCGGCTCTCTACCGCGGCGGCGGCGGCGGCTTCCCGCGGCGGCGGCTCAG	12111
FT	FT	/product= "EvrH"	complement (25230..25233)	FT	FT	406	TTTCG	409
FT	RBS	/**tag= ai	complement (25550..26626)	FT	FT	12112	CCCG	12115
FT	CDS	25550..26626	/**tag= ag	FT	FT			
FT	FT	26685..30479	/product= "EvrI"	FT	FT			
FT	CDS	/**tag= ah	complement (30557..31876)	FT	FT			
FT	FT	26672..26676	/product= "EvrJ"	FT	FT			
FT	RBS	/**tag= ai	complement (31885..31888)	FT	FT			
FT	CDS	/**tag= aj	complement (31941..32882)	FT	FT			
FT	FT	complement (30557..31876)	complement (33167..34405)	FT	FT			
FT	CDS	/**tag= aj	complement (34414..34418)	FT	FT			
FT	FT	complement (31885..31888)	complement (34449..35210)	FT	FT			
FT	CDS	/**tag= ak	complement (35219..35221)	FT	FT			
FT	FT	complement (31941..32882)	complement (35294..36238)	FT	FT			
FT	CDS	/**tag= al	complement (36235..36963)	FT	FT			
FT	FT	complement (30557..31876)	complement (36998..38026)	FT	FT			
FT	RBS	/**tag= am	complement (38072..38566)	FT	FT			
FT	FT	complement (33167..34405)	complement (38892..40163)	FT	FT			
FT	CDS	/**tag= an	complement (40216..40890)	FT	FT			
FT	FT	complement (34414..34418)	complement (40899..40902)	FT	FT			
FT	RBS	/**tag= ao	complement (40887..41576)	FT	FT			
FT	CDS	/**tag= ap	complement (41679..42707)	FT	FT			
FT	FT	complement (34449..35210)	complement (42714..42717)	FT	FT			
FT	RBS	/**tag= ar	complement (42810..43799)	FT	FT			
FT	FT	complement (35219..35221)		FT	FT			
FT	CDS	/**tag= ar		FT	FT			
FT	FT	complement (35294..36238)		FT	FT			
FT	CDS	/**tag= as		FT	FT			
FT	FT	complement (36235..36963)		FT	FT			
FT	RBS	/**tag= at		FT	FT			
FT	FT	complement (36998..38026)		FT	FT			
FT	CDS	/**tag= au		FT	FT			
FT	FT	complement (38072..38566)		FT	FT			
FT	RBS	/**tag= av		FT	FT			
FT	FT	complement (40216..40890)		FT	FT			
FT	CDS	/**tag= av		FT	FT			
FT	FT	complement (40899..40902)		FT	FT			
FT	RBS	/**tag= aw		FT	FT			
FT	FT	complement (40887..41576)		FT	FT			
FT	CDS	/**tag= ax		FT	FT			
FT	FT	complement (41679..42707)		FT	FT			
FT	RBS	/**tag= ay		FT	FT			
FT	FT	complement (42714..42717)		FT	FT			
FT	CDS	/**tag= az		FT	FT			
FT	FT	complement (42810..43799)		FT	FT			

Query Match 9.2%; Score 52.8; DB 22; Length 109519;
Best Local Similarity 48.4%; Pred. No. 0.23;
Matches 147; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

RESULT 13

AA53491/c
ID AAX53491 standard; DNA; 114955 BP.XX AC AAX53491;
XX DT 05-JUL-1999 (first entry)
XX DE Human adenosine A1 receptor antisense oligonucleotide fragment.

XX KW Antisense oligonucleotide; multiple target; antisense treatment;
XX KW impaired respiration; inflammation; lung disease;
XX KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
XX KW acute asthma; allergy; asthma; impeded respiration;
XX KW respiratory distress syndrome; pain; cystic fibrosis;
XX KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
XX KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
XX KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
XX KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
XX KW prostate cancer; ss.

XX OS Synthetic.

XX PN WO9913886-A1.

XX PD 25-MAR-1999.

XX PF 17-SEP-1998; 98WO-US19419.

XX PR 09-JUN-1998; 98US-0093972.

XX PR 17-SEP-1997; 97US-0059160.

XX PA (UYEC-) UNIV EAST CAROLINA.

XX PI Nyce JW;
XX PI WPI; 1999-229400/19.

XX DR New antisense oligonucleotides used in treatment of, e.g. pulmonary
XX PT vasoconstriction
XX PS Disclosure; Page 37; 120pp; English.

XX CC The specification describes antisense oligonucleotides (AAX52869-X55271)
XX CC directed against at least 2 mRNAs selected from target genes, coding and
XX CC non-coding regions of RNAs corresponding to target genes, gene
XX CC initiation codons, genomic flanking regions, intron-exon borders, the
XX CC 5'-end, the 3'-end and the juxta-section between coding and non-coding
XX CC regions and all segments of RNAs encoding proteins associated with one
XX CC or more diseases, conditions or mixtures. The antisense oligonucleotides
XX CC may be derived from sequences AAX5272-74. These multiple target
XX CC oligonucleotides (specifically AAX55180-271) can be used for the
XX CC antisense treatment of diseases and conditions. Typical diseases and
XX CC conditions are those associated with impaired respiration and
XX CC inflammation, including lung diseases, pulmonary vasoconstriction,
XX CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
XX CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
XX CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
XX CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
XX CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
XX CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
XX CC hepatic metastases, as well as all types of cancers which may metastasize
XX CC or have metastasized to the lungs, including breast and prostate cancer.

XX SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 9.2%; Score 52.8; DB 20; Length 114955;
Best Local Similarity 35.6%; Pred. No. 0.24;
Matches 144; Conservative 29; Mismatches 232; Indels 0; Gaps 0;

XX 1 ACCACGGTCCCGCCACGGTCCGACAGCAATCCACACAGCACTTCGACGTCACAG 60

Db 103487 ACGGCCGCGCNCNNNNNSGCCGCAVAGACGGCGCGCNCNNNNNSGCCGCAVAGACGGCGG 103428
Qy 61 GGCGCTGGCAGACACACCAAGCGTGGCACCACCAATGCTTACTACCAAGAGGTGGACT 120
Db 103427 CGCNCNNNNNSACGCGCGCAVAGACGGCGCGCNCNNNNNSGACGCCCAVAGACGGCGCGC 103368
Qy 121 ACTGCTCGGAGGAGGTGAGGTGCGTGGCGCGCGCGCGCTTCGGCGCGCCACGCGCGCGCGG 180
Db 103367 NNHHNNNSGACGCCGCAVAGACGGCGCGCNCNNNNNSCGGAGCGCGCAVAGACGGCGCGG 103308
Qy 181 TCCACGACGACGTCGTCGAAGGAGAGTTCGAGGAGGTGCACACGGTCTTCACGCGCGCGG 240
Db 103307 CNNHHNNNSCGGCGCGCAVAGACGGCGCGCNCNNNNNSGCCGCGCGCAVAGACGG 103248
Qy 241 CCAACCCACACACACCATGTGTACACGCGCGCGCGCGCTTCGTGTGCGCGGAGACCA 300
Db 103247 CGCGCNCNNNNNSGCCGCGCGCAVAGACGGCGCGCNCNNNNNSGCCGCGCGCGG 103188
Qy 301 GGGTCGAAGAGACATCAACACCTGCACGCGCGAGGTCCACGAGCGCGAGGAGAGCTTCC 360
Db 103187 CCAVAGACGGCGCGCNCNNNNNSCGGCGCGCGCAVAGACGGCGCGCNCNNNNNSCC 103128
Qy 361 TCGCCAGGCGTAACTGACGCGCGCGCGCGCGCGCATCCACGCGCGG 405
Db 103127 CGGCGCGGCGCGCGCAVAGACGGCGCGCNCNNNNNSGCCGCGCGCGG 103083

RESULT 14

AAT10429

ID AAT10429 standard; DNA; 2244 BP.

XX AC AAT10429;
XX DT 03-JUN-1996 (first entry)

XX DE Mature isoamylase DNA.
XX KW Isoamylase; starch; amylose; amylopectin; transgenic plant;
XX KW fructose; ds.
XX OS Flavobacterium sp.

XX PN WO9603513-A2.
XX PD 08-FEB-1996.
XX PF 24-JUL-1995; 95WO-US09323.
XX PR 28-JUL-1994; 94US-0281902.

XX PA (MONS) MONSANTO CO.
XX PI Barry GF, Kishore GM, Krohn BW;
XX DR WPI; 1996-117056/12.
XX DR P-PSDB; AAR92693.

XX PT New isolated DNA encoding isoamylase - used to produce plants having
XX PT a modified starch content, e.g. a higher ratio of amylose to
XX PT amylopectin
XX PS Claim 2; Page 35-39; 46pp; English.

XX CC A DNA sequence (AAT10429) codes for the mature isoamylase (plus an
XX CC N-terminal methionine) of Flavobacterium sp. (see also AAR92692).
XX CC It was obtd. by amplification of the cloned isoamylase gene
XX CC (AAT10428) using primers (AAT10432-33) designed to add a start codon
XX CC to the beginning of the processed mature enzyme (Ala-33) and a
XX CC stop codon adjacent to the endogenous TGA codon. The DNA sequence
XX CC is used for the recombinant prodn. of isoamylase or can be
XX CC expressed in transgenic plants, e.g. potato, cassava, sweet potato,
XX CC corn, wheat, barley or rice, to alter the amylose:amylopectin

CC ratio of the starch.

XX Sequence 2244 BP; 427 A; 827 C; 704 G; 286 T; 0 other;

Query Match 9.08; Score 51.8; DB 17; Length 2244;

Best Local Similarity 54.5%; Pred. No. 0.25;

Matches 104; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 151 CGCGCGCTTCGGCGCCACGCGCGGCGTCCAGCAGCAGCTCGTCAAGGAGAAGTTCCG 210

DB 1877 CGGCCAGGACACCAACGGCAACGTGATGAGCAGTTGCGTGGTTCAAGCCCGACGGCG 1936

QY 211 AGGAGGTCGACACGGTCTCAGCGCGCGGCCCAACCAACCACCATGGTCCACACG 270

DB 1937 CGCAGGCGCAGCGCGCTACTTTCAACGCGCGGCAACCAACCGCCCTGGCGGCATCG 1996

QY 271 GCGGCCACGGCTTCGTGGTGGCGGAGACCAAGAGGTCGAAGAGGACATCAACACCTGCACCG 330

DB 1997 ACGCCACGAGTTTCGGCGACAGCGCCAGCGCGATCTACGTGCGCTACAACGGCTGGTCCG 2056

QY 331 GCGAGGTCCAC 341

DB 2057 GCGCGGTCCGAC 2067

AC AAV23640;

XX 17-JUL-1998 (first entry)

XX Flavobacterium isoamylase coding sequence.

XX Isoamylase; high amylose starch production; ds.

XX Flavobacterium sp.

XX Key Location/Qualifiers

XX CDS 1..2241

XX /*tag= a

XX US5750876-A.

XX 12-MAY-1998.

XX 07-JUN-1995; 95US-0476519.

XX 07-JUN-1995; 95US-0476519.

XX 28-JUL-1994; 94US-0281902.

XX (MONS) MONSANTO CO.

XX Barry GF, Kishore GM, Krohn BM;

XX WPI: 1998-296898/26.

XX P-PSDB: AAW53886.

XX DNA coding for Flavobacterium isoamylase - for production of

XX recombinant isoamylase or transgenic plants

XX Claim 2; Column 27-32; 20pp; English.

XX This sequence encodes the Flavobacterium isoamylase of the invention. The

XX protein can be used to produce higher amylose starches, using genetically

XX modified plants.

XX Sequence 2244 BP; 427 A; 828 C; 703 G; 286 T; 0 other;

Query Match 9.08; Score 51.8; DB 19; Length 2244;

Best Local Similarity 54.5%; Pred. No. 0.25;

Matches 104; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 151 CGCGCGCTTCGGCGCCACGCGCGGCGTCCAGCAGCAGCTCGTCAAGGAGAAGTTCCG 210

DB 1877 CGGCCAGGACACCAACGGCAACGTGATGAGCAGTTGCGTGGTTCAAGCCCGACGGCG 1936

QY 211 AGGAGGTCGACACGGTCTCAGCGCGCGGCCCAACCAACCACCATGGTCCACACG 270

DB 1937 CGCAGGCGCAGCGCGCTACTTTCAACGCGCGGCAACCAACCGCCCTGGCGGCATCG 1996

QY 271 GCGGCCACGGCTTCGTGGTGGCGGAGACCAAGAGGTCGAAGAGGACATCAACACCTGCACCG 330

DB 1997 ACGCCACGAGTTTCGGCGACAGCGCCAGCGCGATCTACGTGCGCTACAACGGCTGGTCCG 2056

QY 331 GCGAGGTCCAC 341

DB 2057 GCGCGGTCCGAC 2067

Search completed: June 16, 2003, 02:09:39

Job time : 214.272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 23:43:51 ; Search time 1584.35 Seconds
(without alignments)
9917.805 Million cell updates/sec

Title: US-10-090-035-3

Perfect score: 574

Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaa 574

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	174.6	30.4	173699	8	AP003416	AP003416 Oryza sat
C 2	174.6	30.4	174218	2	AC091071	AC091071 Oryza sat
C 3	174.6	30.4	185095	8	AP003436	AP003436 Oryza sat
C 4	174.6	30.4	185481	2	AC091088	AC091088 Oryza sat
C 5	143.8	25.1	562	8	SST242802	AT242802 Sporeobol
C 6	62.6	10.9	390	6	AR135147	AR135147 Sequence
C 7	61.4	10.7	35100	1	SC3A7	AL031155 Streptomy
C 8	61.4	10.7	150676	2	AC107225	AC107225 Oryza sat
C 9	58.2	10.1	29080	1	SCBAC28G1	AL593842 Streptomy
C 10	57.8	10.1	18420	1	SCC30	AL593972 Streptomy
C 11	57	9.9	4851	1	SC018817	Y18817 Streptomyce
C 12	56.8	9.9	150053	8	AP003735	AP003735 Oryza sat
C 13	56.2	9.8	102688	2	AP004045	AP004045 Oryza sat
C 14	56.2	9.8	138289	2	AP004873	AP004873 Oryza sat
C 15	54.6	9.5	3164	1	AF263463	AF263463 Streptomy
C 16	54.4	9.5	13368	1	AE005107	AE005107 Halobacte
C 17	54.2	9.4	106050	2	AP004088	AP004088 Oryza sat
C 18	54.2	9.4	146436	2	AC108759	AC108759 Oryza sat
C 19	54.2	9.4	152263	2	AP005412	AP005412 Oryza sat
C 20	54.2	9.4	187705	2	AP004765	AP004765 Oryza sat
C 21	54	9.4	36368	1	SC985	AL035206 Streptomy
C 22	53.8	9.4	174410	2	AC128158	AC128158 Rattus no
C 23	53.6	9.3	34593	1	SC2K8	AL589164 Streptomy
C 24	53.4	9.3	10612	1	AF086832	AF086832 Streptomy
C 25	53.4	9.3	41173	1	SCI41	AL132648 Streptomy
C 26	53.4	9.3	65547	8	AY013245	AY013245 Oryza sat
C 27	53.4	9.3	120159	2	AC120509	AC120509 Oryza sat
C 28	53.2	9.3	15041	1	AE006991	AE006991 Mycobacte
C 29	53.2	9.3	47852	1	MTV023	AL022022 Mycobacte
C 30	53	9.2	136150	8	AP002485	AP002485 Oryza sat
C 31	52.8	9.2	1128	6	AX393069	AX393069 Sequence
C 32	52.8	9.2	1128	9	AB040799	AB040799 Homo sapi
C 33	52.8	9.2	1245	6	AX195950	AX195950 Sequence
C 34	52.8	9.2	109519	6	AX195929	AX195929 Sequence
C 35	52.8	9.2	163494	2	AC016964	AC016964 Homo sapi
C 36	52.8	9.2	171297	9	AC096970	AC096970 Homo sapi
C 37	52.8	9.2	190506	2	AC068990	AC068990 Homo sapi
C 38	52.6	9.2	179198	2	AC130729	AC130729 Oryza sat
C 39	52	9.1	41622	1	SCD25	AL118514 Streptomy
C 40	52	9.1	143676	2	AP005182	AP005182 Oryza sat
C 41	51.8	9.0	2244	6	AR007590	AR007590 Sequence
C 42	51.8	9.0	2334	6	AR007585	AR007585 Sequence
C 43	51.8	9.0	2336	1	FS090120	U90120 Flavobacter
C 44	51.8	9.0	38532	1	SCE94	AL049628 Streptomy
C 45	51.8	9.0	149256	2	AP003827	AP003827 Oryza sat

ALIGNMENTS

RESULT 1

AP003416/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AP003416
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0456E05.
AP003416
AP003416.4 GI: 20804922
Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,
clone: P0456E05.
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

CDS

```
complement(join(40277..40417,40717..40773,40875..41073,
41173..41300,41399..41504,41677..41756,42145..42242,
42324..42437,42523..42663,42757..42841,43717..44055))
/gene="P0456E05.7"
/notes="contains ESTs AU101665(R2414),D24709(R2594)
similar to Arabidopsis thaliana chromosome 5, At5g14420
unknown protein"
/codon_start=1
/protein_id="BAB92608.1"
/db_xref="GI:20804929"
/translation="MGGGSRSPDRSGHGRYCHSPFQQQWGGGGGGGGGGGYPY
GDAHHGVGYAPAGQGYAAYPAYQPPMPSPAAQTPPARAGGASKPLDRYS
RADDDHYVDQVTDALAQGLESSLNIVGIDFTKSNWLGKSFHGRSLHLSNAPNE
YEQAISIGTILSKFDEDNLPCFGGDATTHDQVCFDPLRCNCFSEALSRYE
LVPLHLRAGTSPFIENAMTIVSQSGQYHVLIIADGQVTRSDVAFSGQLSSQEQ
KQSDATVASELPLSLVILVGVGDGDMKFEEDNIPSAFDFNFVAFSEIMKNMP
QSRKEAFAALSMALPEIQYKATVGLGILGRSFSPERVPLPPGGSHDAYSYGSKS
FKPSTYPOSSYSSSYPHYETAQSSPAVPSSTYDNOVPCILVNPDKMAFGCGHQ
NPKAFERLADRPVYKWCIDRKASNPQLQDCGAGDVSSIETH"
51367..52887
/gene="P0456E05.8"
51367..52887
/gene="P0456E05.8"
/notes="hypothesized protein
similar to salt-inducible protein"
/codon_start=1
/protein_id="BAB92609.1"
/db_xref="GI:20804930"
/translation="NAASRRLLKPLSLISKHORLISPEIDVEQAASPASSSIPLD
PSLPLAVSHLSPPLPALPSVHASTPAALRLRRARHHPRLAALDHLHLLAA
SDSPAFRDLSTLSALASRLRLSLRLLELVLPFCPCDSDISFACPELLTF
RAILAFADGDIIPAASDALASRLRAADSPFABFYNIILHALARRHEDTVRYGE
MTSHRVHPDYVTINILINSVRAEVDAAWFOEMRWSCSPGVSFTLMGFER
ERCKEGLKVARMLQGLFGLSVASMEIMINGLRCGGEPKAAEVFEFLVDGVVPG
FDCLDVESLCVRNVEKAVEVELILERNVSLGVPAGTVVLECLMKGLDKACQ
MNGRWAAEIVPDITSCNYIFEALCEAGRTVDANRLQAKERGFQADGFTYSIVGG
FGRQIRKEGAVALDEMDAGYVFNATYINRLDLGHMSMOIQKCSRHNAN"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
complement(join(54710..55150,55176..55310,55441..55668))
/gene="P0456E05.9"
```

Query Match

```
Best Local Similarity 30.4%; Score 174.6; DB 8; Length 173699;
Matches 257; Conservative 0; Mismatches 84; Indels 12; Gaps 2;

QY 76 CACACCAAGCGTCGGCACCACATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAG 135
DB 157591 CACTCCCCAACATAAAACAAATGGCTCACTACACGAGGTGGACTACTGCTCGGAGGAG 157532

QY 136 TGAGTCGGTGGCCCC---CGGCCGGCTTCGCCGCCACGCGCGGGCGTCCAGCAGCAG 192
DB 157531 TGAGTCGGTGGACCCCGCCCGCGGCTTCTTCGGCGCGCGCGGCGTCCAGCAGCAG 157472

QY 193 TCCTCAAGGAGAGTTCGAGGAGGTTCGACACGGTCTCACCGCGCGCGGCCAACCCACC 252
DB 157471 TCCTCAAGGAGAGCTTCGAGGAGGTTCGACACGGTCTCACCGCGCGGCCAACCCACC 252

QY 253 ACCACCATGGTCACACCGCGCGGACCGGCTTCGTGGTGCAGAGACCAAGGTGGAGGAG 312
DB 157420 ACCACCAACCAACACCGGCAACGACTACCTGATGGTGCAGAGACCAAGGTGGAGGAG 157361

QY 313 ACATCAACACCTGCACCGCGCGGACCGGCTTCGTGGTGCAGAGACCAAGGTGGAGGAG 372
DB 157360 ACTTCAACACCTGCACCGCGCGGAGTTCGGGAGCGCAGCAGAGCTTCCCTCAAGTCG 157301

QY 373 ACTGAGCGCGCGCGCGCGGACATCACGCCCTTCGTGGTTCCTTCGCTGCCGTGCC 425
DB 157300 ACTGATCGAACCCTGCTGTGTGTACCCGCTGTACGTACGTATATGTGTGCC 157248
```

RESULT 2
AC091071

```
LOCUS
DEFINITION
  Oryza sativa (japonica cultivar-group) DNA linear HTG 27-MAR-2001
  *** SEQUENCING IN PROGRESS ***, 3 ordered pieces.
ACCESSION
  AC091071
VERSION
  HTG: HTGS_PHASE2.
SOURCE
  Oryza sativa (japonica cultivar-group).
  Oryza sativa (japonica cultivar-group).
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
  1 (bases 1 to 174218)
  Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Lee, M.-C., Yun, D.-W., Kim, H.-I.
  and Eun, M.Y.
  Oryza sativa PAC P0456E05 genomic sequence
  Unpublished
  2 (bases 1 to 174218)
  Hahn, J.-H., Eun, M.Y. and Kim, H.-I.
  Direct Submission
  Submitted (27-MAR-2001) Rice Genome Sequencing Project, National
  Institute of Agricultural Science and Technology(NIAST), RDA, 249
  Seodun-dong, Suwon 441-707, Korea (E-mail:jhahn@rda.go.kr,
  Tel:82-31-290-0309, Fax:82-31-290-0308)
  The PAC clone was provided by Japanese Rice Genome Research Program
  of NIAST.
```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 53113: contig of 53113 bp in length
* 53114 53213: gap of unknown length
* 53214 172404: contig of 119191 bp in length
* 172405 172504: gap of unknown length
* 172505 174218: contig of 1714 bp in length.
* Location/Qualifiers
  1..174218
    /organism="Oryza sativa (japonica cultivar-group)"
    /variety="Nipponbare"
    /db_xref="taxon:39947"
    /chromosome="1"
    /clone="P0456E05"
```

BASE COUNT 49037 a 38377 c 37185 g 49417 t 202 others
ORIGIN

```
Query Match 30.4%; Score 174.6; DB 2; Length 174218;
Best Local Similarity 72.8%; Pred. No. 5.4e-17;
Matches 257; Conservative 0; Mismatches 84; Indels 12; Gaps 2;

QY 76 CACACCAAGCGTCGGCACCACATGGCTTACTACACGAGGTGGACTACTGCTCGGAGGAG 135
DB 16041 CACTCCCCAACATAAAACAAATGGCTCACTACACGAGGTGGACTACTGCTCGGAGGAG 16100

QY 136 TGAGTCGGTGGCCCC---CGGCCGGCTTCGCCGCCACGCGCGGGCGTCCAGCAGCAG 192
DB 16101 TGAGTCGGTGGACCCCGCCCGGCTTCTTCGGCGCGCGGCGGCTTCAGCAGCAGCAG 16160

QY 193 TCCTCAAGGAGAGTTCGAGGAGGTTCGACACGGTCTCACCGCGCGCGGCCAACCCACC 252
DB 16161 TCCTCAAGGAGAGCTTCGAGGAGGTTCGACACGGTCTCACCGCGCGGCCAACCCACC 16211

QY 253 ACCACCATGGTCACACCGCGCGGACCGGCTTCGTGGTGCAGAGACCAAGGTGGAGGAG 312
DB 16212 ACCACCAACCAACACCGGCAACGACTACCTGATGGTGCAGAGACCAAGGTGGAGGAG 16271

QY 313 ACATCAACACCTGCACCGCGCGGACCGGCTTCGTGGTTCCTTCGCTGCCGTGCC 372
DB 16272 ACTTCAACACCTGCACCGCGCGGAGTTCGCGGAGCGCAGAGCTTCTTCTCAAGTCG 16331
```

Mon Jun 16 14:55:56 2003

QY	373	ACTGAGCGCCGCGGCGGCATCCAGCCGCTGCGTGGTGGCTGCGTCC	425
Db	16332	ACTGATCGACCTGCTGTGTACCGCGTGTACGTACGTATGTGTGCC	16384
RESULT 3			
AP003436/c			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,		
ACCESSION	AP003436		
VERSION	AP003436.2	GI:16197551	
KEYWORDS			
SOURCE			
ORGANISM	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,		
	clone:P0470A12.		
	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC		
	clone:P0470A12		
JOURNAL	Published Only in Database (2001)		
REFERENCE	2	(bases 1 to 185095)	
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAR-2001) Takuji Sasaki, National Institute of		
	Agrobiological Sciences, Rice Genome Research Program; Kamondai		
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
	(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,		
	Tel:81-298-38-7441, fax:81-298-38-7468)		
COMMENT	On Oct 17, 2001 this sequence version replaced gi:13430001. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from SP6 to P7 of the PAC clone. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.		
FEATURES	Location/Qualifiers		
source	1..185095		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:39947"		
	/chromosome="1"		
	/clone="P0470A12"		
gene	join(10813..10841,11094..11205,11289..11432,11817..11921,		
	12046..12129,12315..12430,12535..12636,12715..12828,		
	12916..12972,13222..13288,13356..13409,13798..13872,		
	13956..13991,14389..14490)		
	/gene="P0470A12.1"		
CDS	join(10813..10841,11094..11205,11289..11432,11817..11921,		
	12046..12129,12315..12430,12535..12636,12715..12828,		
	12916..12972,13222..13288,13356..13409,13798..13872,		
	13956..13991,14389..14490)		
	/gene="P0470A12.1"		
	/note="hypothetical protein		
	similar to Oryza sativa chromosome 5, P0574H01.5"		
	/codon_start=1		
	/protein_id="BAB90276.1"		
	/db_xref="GI:20161352"		
	/translation="MSKLAIGRSSGLPRSNMGRLLFSAVIGVLMGLYLGISFTVN		
	VTKLHPSSIIYIEDKDSGITQTLLNHAWTSANSKRNNSSESDIKPIYVPTNP		
	KGAELAPGIVPETDIYLRLLGEPSEDLTSPRYLITFTVYISKANIDAAVKRFS		
	ENFTIMLHYDGRINDWDEFWSKRAIHVSVRQTKWAKREFLHPDIAVPDYIFIW		
	DELSVQHNAEAYIKLVRRKHLEISQPLGPKGLTWQTKRLGLDQEVKHYTEERP		
	WCTDPLHPCAAEIATVSRDAWRVVMHMTQNDLVHGWGLDFALRCRCVFAHEKI		
	GVDSQWVTHOIVPSLNGNQTAEGRTPWEGVRARCKEWMGFQKRLADAKEYILGK		
	GITPPN"		
	complement(15118..16344)		
	/gene="P0470A12.2"		
	complement(15118..16344)		
	/gene="P0470A12.2"		
	/note="contains EST AU069076(c51993)		
	unknown protein"		
	/codon_start=1		
	/protein_id="BAB90277.1"		
	/db_xref="GI:20161353"		
	/translation="MAMRNGNGYQYAYSYGYSKPOVNYHSOSSSVTVVTIKM		
	DEMSVHENAVKMSSSFREKRYEQGGAGDVQVCRNGGAAGAVQKHVYEKKKEE		
	VYESNHGGGGCGGYARHGAGGQKHATYQHEVLKYESAGGTAARHNGAGAG		
	GVKKOYSYREEEYDAGYARHNGGYQKQYSSYRDEEDCAAAGGYYGRHNGAGG		
	AVKOHATYKQHAITEGVNGCAGGYNNRHQAVAGAGGHHYGGATAAYANASNKQ		
	HFTAAAGHHSSGGHTOYHHQSYECEDEDEDDDDDDDDDDGDCPPSRQC		
	SVSHYQHAQHEKQKAGGRNHAYERHEEHGGGAQRYQKYESSTQVYAGGGGN		
	GGCYRGNAPCFRIMA"		
	complement(join(19564..19724,21382..21688,22059..22217,		
	22387..22505,22923..23035,23239..23318))		
	/gene="P0470A12.3"		
	complement(join(19564..19724,21382..21688,22059..22217,		
	22387..22505,22923..23035,23239..23318))		
	/gene="P0470A12.3"		
	/note="hypothetical protein"		
	/codon_start=1		
	/protein_id="BAB90278.1"		
	/db_xref="GI:20161354"		
	/translation="MARRVGGTSEVGGNGEARGRRQRSGRREARGARRRLPRRG		
	SPASPPSRBENGEGSGGRAPVINTVTRSRKQWRDSSGRRDHAASRHDSIHT		
	REKTWNAGIKLPIFTVTRFCRCPSLVPTPWAGILGCHLVRLALCLLRQWR		
	RHAAHAVRNALGVVVRVLPVPAATMRKQCATGAALRYEAHTTENTTIGE		
	TDDQATVTRNTLKRRTLCMSVCHRGVQVQVQAPYQG"		
	complement(23929..24237)		
	/gene="P0470A12.4"		
	complement(23929..24237)		
	/gene="P0470A12.4"		
	/note="hypothetical protein"		
	/codon_start=1		
	/protein_id="BAB90279.1"		
	/db_xref="GI:20161355"		
	/translation="MAQYSEVDHCAEEMRPPHAGGEHYAVRRESEYEVDEMARAGR		
	GHHNGGGGGHLYSGSRHGDALHGHREELVHGHGHRGGGRQYDSTGYYG		
	"		
	complement(join(27239..27499,30542..30622,30720..30899,		
	30998..31165))		
	/gene="P0470A12.5"		
	complement(join(27239..27499,30542..30622,30720..30899,		
	30998..31165))		
	/gene="P0470A12.5"		
	/note="contains EST AU070442(S15668)		
	unknown protein"		
	/codon_start=1		
	/protein_id="BAB90280.1"		
	/db_xref="GI:20161356"		
	/translation="MDSGGTASPVAGDAGGRSVFRGDAPKLALAKEMKEGLD		
	LVTKVKALTRKKNQLPTADGIGVLEAKHLLISYCODIVYLLRKAAGLSVEHPG		
	VYRSLEIFRLFLKPIEDIKKMEYQIKLTNRADSCAQAQEKVDYSEVRSVTPGPG		
	FLGRGGVQOOHVKEITFQIDRSRGRHHNNHNDYLMVRETKVEEDFNCTGEFR		
	ERKQSFLKSD"		
	complement(join(34470..34734,34836..34921,35025..35117,		
	35201..35364,35440..35601,35682..35861,35978..36414,		
	36513..36568))		

Mon Jun 16 14:55:56 2003

us-10-090-035-3_1.rge

[illegible]

RESULT 5				
SST242802				
LOCUS	SST242802	562 bp	mRNA	linear PLN 27-NOV-1999
DEFINITION	<i>Scorobolus stansianus</i> mRNA for putative glycine-rich protein.			

DEFINITION	01	Accession	AJ242802
ACCESSION	02	Version	AJ242802.1
VERSION	03	Keywords	GI:6478143
KEYWORDS	04	Source	glycine-rich protein.
SOURCE	05	Organism	Sporobolus stapfianus.
ORGANISM	06		Sporobolus stapfianus.
	07		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
	08		clade; Chloridoideae; Eragrostideae; Sporobolus.
REFERENCE	09		1 (bases 1 to 562)
AUTHORS	10		Neale,A.D., Blomstedt,C.K., Bronson,P., Le,T.N., Guthridge,K., Evans,J.J., Gaff,D.F. and Hamill,J.D.
TITLE	11		The isolation of lowly-transcribed genes which are induced during dedensation of the resurrection grass Sporobolus stapfianus
JOURNAL	12		Unpublished
REFERENCE	13		2 (bases 1 to 562)
AUTHORS	14		Neale,A.D.
TITLE	15		Direct Submission
JOURNAL	16		Submitted (27-MAY-1999) Neale A.D., Biological Sciences, Monash University, Wellington Rd Clayton, Victoria, 3168, AUSTRALIA
FEATURES	17		Location/Qualifiers
source	18		1..562

```

/organism="Sporobolus stapfianus"
/db_xref="taxon:56623"
/clone="SDG137c"
63..350
/codon_start=1
/product="putative glycine-rich protein"
/protein_id="CAB61838.1"
/db_xref="GI:6478144"
/translation="MAHFEMVDYEVTSMKPGRGGGGVQGVVVKETFEVEQV
TPRSGHHGGHGRNGCHSGHQARETFERHQHAHRVPRQGERPCQG"
143 a 153 c 154 g 112 t
BASE COUNT
ORIGIN

```

	Query Match	25.1%	Score 143.8	DB 8	Length 562
	Best Local Similarity	62.1%	pred. No. 4e-12		
	Matches 321	Conservative 0	Mismatches 172	Indels 24	Gaps 5
QY	79	ACCAAGCTCGGCACCAATGGCTTACTATACAGAGAGGTGACACTACTCTCGGAGGAGGTCA	138		
Db	46	ACCAATCAAGAATACAAATGCGGCACCTTCAAGGAGATGGTGGACTACGAGGAGGTGACGT	105		
QY	139	GTGTCGGTGCCTCGCGCGGCTTTCGCCCGCCACGCGCGCGCTTCACAGCAGCAGTGTGTC	198		

106	CGATGGCCAAACGGGCTTCGCGCCACACGGCGGGCGGCGTTCAGACAGTTCTGCTGTCA	165
Db		
199	AGGAGAGTTCGAGGAGGTTCAGACGGTCTCACGCGCGCGGCCAACACACACCA	253
Qy		
166	AGGAGACGTTTCGAGGAGGTTCGAGCAGGTACCCACGTTGGCGCAGCGCCACACGCGT	225
Db		
254	-----CCACCATGTCTACCACGGCGGCACGGGTTCGTGTGTCGCGAGACCGAGGTCTG	306
Qy		
226	GCCACCGCGGCACCAAGCCACGGCTCGCGGCACCTTCAGGCGCGCGAGACCAAGTTTG	285
Db		
307	AAGAGGACATCAACACCTTCGACCGGGGAGGTTCACAGCGCGCAGGAGAGCTTCCTTCGCCA	366
Qy		
286	A--GAGACATCAACACGCGCCGCGGAGTTCACGAGCGCAAGGAGACAGCTCTTGTC	343
Db		
367	GGGCTRACTAGCCGCGCGCGCGGCATCCACGCCCGTTCGTGCTTCGCTCGCTGGCT	426
Qy		
344	GGGCTGACTGAGCAGCTTCGTGCTCTGCTGCACGTTTCGTATGTAACCTGTGTGAAGG	403
Db		
427	TATGTATGTCGTGGTTGACTGGTTGT-----TCAGGGTCATCGTACTTTGGCTATCTGC	481
Qy		
404	TCATAAGTCGTACCCGATGTATGTATGTGTGTGTCGCCGTACTTTGGCTGCAGCTGC	463
Db		
482	GTGCACGCATCTAGCTCCTCTAGCAATTCAGCAATAAGCTCGTGCACCTG---AATAAA	537
Qy		
464	ATGCACGCACGGCGCTTCTGTACGAACTACCA--AATAATTCATGACCTGAATAATAA	522
Db		
538	ACTTCTTCGTAATTAATAAAAAAAAAAAAAAAAAAAAAA	574
Qy		
523	GCATTTCCGTATGTTGTTGTATATAAAAAAAAAAAAAA	559
Db		

RESULT 6

AR135147

LOCUS AR135147 390 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6194550.

ACCESSION AR135147

VERSION AR135147.1 GI:14124052

RESULT	6
ARI35147	
LOCUS	ARI35147
DEFINITION	Sequence 7 from patent US 6194550.
ACCESSION	ARI35147
VERSION	ARI35147.1 GI:14124052
	390 bp DNA
	linear
	PAT 16-MAY-2001

KEYWORDS	.	Unknown.	
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 390)	
AUTHORS		Gold, L., Tuerk, C., Pribnow, D. and Smith, J. Drew.	
TITLE		Systematic polypeptide evolution by reverse translation	
JOURNAL		Patent: US 6194550-A 7 27-FEB-2001;	
FEATURES		Location/Qualifiers	
	source	1..390	
		/organism="unknown"	
BASE COUNT		125 a	126 c
ORIGIN			133 g
			6 t

Query Match	10.9%;	Score 62.6;	DB 6;	Length 390;
Best Local Similarity	52.0%;	Pred. No. 4.1;		
Matches 140;	Conservative 0;	Mismatches 129;	Indels 0;	Gaps 0;
QY	103	ACTACAGAGGTGACTACTGCTCGGAGGAGGTGAGTGGTCGGTCGGCCCCGGCCGGCTTCG	162	
Db	109	ACGAGCAGCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	168	
QY	163	GCCGCCACGGCGCGCGTTCAGCAGCACGTCGTCACAGGAAAGTTCGAGGAGGTCGACA	222	
Db	169	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	228	
QY	223	CGGTCTCACGCGCCGGCGGCGCAACACACACACCCATGTTCAACCGGCGGCCACCGCT	282	
Db	229	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	288	
QY	283	TCGTGTGTCGCGAGACCCAGGGTCGAAGAGGAGCATCAACACCTGCACCGCGGAGGTCACG	342	
Db	289	ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	348	
QY	343	AGCGCAGGGGAGGACTTCCTTCGCCGACGGCT	371	

[illegible]

BASE COUNT	41031 a	35084 c	34425 g	40132 t	4 others
ORIGIN	/clone="OSJNBa0026H19"				
Query Match	10.7%; Score 61.4; DB 2; Length 150676;				
Best Local Similarity	51.3%; Pred. No. 2.8;				
Matches 143;	Conservative 0; Mismatches 136; Indels 0; Gaps 0;				
QY	110	GGAGGTGGACTACTGCTCGGAGGAGGTGAGTTCGTGGCCCGCGCGCTTCGGCCGCCCA	169		
Db	55347	GGAGGTGGTGTACCCGAAGGCGCATGTCGACCCGGACGCCAACAGGGCTTCGAACCT	55288		
QY	170	CGCGCGCGCGTTCGAGCAGCAGCTCGTCAAGAGAGAAAGTTCGAGGAGGTTCGACAGGTCTC	229		
Db	55287	CCGGCTCTGCCCTATCTGACAGCGCGCCGCGCAGCCGAGGCGCTCTCGCGGGGT	55228		
QY	230	AGCGCCGCGCGCAACACACACACACACACCTGTCACCGCGCGAGGTTCGTGGGT	289		
Db	55227	GCTCGCGGGAGGTACCAACGTCGCGGACGACACCGGCGCCCGCACACCGCGCGCA	55168		
QY	290	GGCGGAGACACAGGTTCGAAGAGGATCAACACCTTCACCGCGCGAGGTTCACGAGCGAG	349		
Db	55167	GATCTCTCCCAAGGTGGAGGAGTTGATGGCCCGCATCACCGCGCAGGTTCGACGCCCGAG	55108		
QY	350	GGAGAGGTTCTCGCCAGGGCTAACTAGCGCGCCGGCG	388		
Db	55107	CACAGCAGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	55069		

RESULT_9	
SCBAC28G1/c	
LOCUS	SCBAC28G1
DEFINITION	Streptomyces coelicolor cosmid BAC28G1.
ACCESSION	ALF53842
VERSION	ALF53842.2
KEYWORDS	actII-1; actII-2; actII-3; actII-4; actII-5;
	actIII-1; actIII-2; actIII-3; actIII-4; actIII-5;
	actIV-1; actIV-2; actIV-3; actIV-4; actIV-5;
	actV-1; actV-2; actV-3; actV-4; actV-5;
	actVI-1; actVI-2; actVI-3; actVI-4; actVI-5;
	actVII-1; actVII-2; actVII-3; actVII-4; actVII-5;
	actVIII-1; actVIII-2; actVIII-3; actVIII-4; actVIII-5;
	actIX-1; actIX-2; actIX-3; actIX-4; actIX-5;
	actX-1; actX-2; actX-3; actX-4; actX-5;
	actXI-1; actXI-2; actXI-3; actXI-4; actXI-5;
	actXII-1; actXII-2; actXII-3; actXII-4; actXII-5;
	actXIII-1; actXIII-2; actXIII-3; actXIII-4; actXIII-5;
	actXIV-1; actXIV-2; actXIV-3; actXIV-4; actXIV-5;
	actXV-1; actXV-2; actXV-3; actXV-4; actXV-5;
	actXVI-1; actXVI-2; actXVI-3; actXVI-4; actXVI-5;
	actXVII-1; actXVII-2; actXVII-3; actXVII-4; actXVII-5;
	actXVIII-1; actXVIII-2; actXVIII-3; actXVIII-4; actXVIII-5;
	actXIX-1; actXIX-2; actXIX-3; actXIX-4; actXIX-5;
	actXX-1; actXX-2; actXX-3; actXX-4; actXX-5;
	actXXI-1; actXXI-2; actXXI-3; actXXI-4; actXXI-5;
	actXXII-1; actXXII-2; actXXII-3; actXXII-4; actXXII-5;
	actXXIII-1; actXXIII-2; actXXIII-3; actXXIII-4; actXXIII-5;
	actXXIV-1; actXXIV-2; actXXIV-3; actXXIV-4; actXXIV-5;
	actXXV-1; actXXV-2; actXXV-3; actXXV-4; actXXV-5;
	actXXVI-1; actXXVI-2; actXXVI-3; actXXVI-4; actXXVI-5;
	actXXVII-1; actXXVII-2; actXXVII-3; actXXVII-4; actXXVII-5;
	actXXVIII-1; actXXVIII-2; actXXVIII-3; actXXVIII-4; actXXVIII-5;
	actXXIX-1; actXXIX-2; actXXIX-3; actXXIX-4; actXXIX-5;
	actXXX-1; actXXX-2; actXXX-3; actXXX-4; actXXX-5;
	actXXXI-1; actXXXI-2; actXXXI-3; actXXXI-4; actXXXI-5;
	actXXXII-1; actXXXII-2; actXXXII-3; actXXXII-4; actXXXII-5;
	actXXXIII-1; actXXXIII-2; actXXXIII-3; actXXXIII-4; actXXXIII-5;
	actXXXIV-1; actXXXIV-2; actXXXIV-3; actXXXIV-4; actXXXIV-5;
	actXXXV-1; actXXXV-2; actXXXV-3; actXXXV-4; actXXXV-5;
	actXXXVI-1; actXXXVI-2; actXXXVI-3; actXXXVI-4; actXXXVI-5;
	actXXXVII-1; actXXXVII-2; actXXXVII-3; actXXXVII-4; actXXXVII-5;
	actXXXVIII-1; actXXXVIII-2; actXXXVIII-3; actXXXVIII-4; actXXXVIII-5;
	actXXXIX-1; actXXXIX-2; actXXXIX-3; actXXXIX-4; actXXXIX-5;
	actXXXX-1; actXXXX-2; actXXXX-3; actXXXX-4; actXXXX-5;
	actXXXXI-1; actXXXXI-2; actXXXXI-3; actXXXXI-4; actXXXXI-5;
	actXXXXII-1; actXXXXII-2; actXXXXII-3; actXXXXII-4; actXXXXII-5;
	actXXXXIII-1; actXXXXIII-2; actXXXXIII-3; actXXXXIII-4; actXXXXIII-5;
	actXXXXIV-1; actXXXXIV-2; actXXXXIV-3; actXXXXIV-4; actXXXXIV-5;
	actXXXXV-1; actXXXXV-2; actXXXXV-3; actXXXXV-4; actXXXXV-5;
	actXXXXVI-1; actXXXXVI-2; actXXXXVI-3; actXXXXVI-4; actXXXXVI-5;
	actXXXXVII-1; actXXXXVII-2; actXXXXVII-3; actXXXXVII-4; actXXXXVII-5;
	actXXXXVIII-1; actXXXXVIII-2; actXXXXVIII-3; actXXXXVIII-4; actXXXXVIII-5;
	actXXXXIX-1; actXXXXIX-2; actXXXXIX-3; actXXXXIX-4; actXXXXIX-5;
	actXXXXX-1; actXXXXX-2; actXXXXX-3; actXXXXX-4; actXXXXX-5;
	actXXXXXI-1; actXXXXXI-2; actXXXXXI-3; actXXXXXI-4; actXXXXXI-5;
	actXXXXXII-1; actXXXXXII-2; actXXXXXII-3; actXXXXXII-4; actXXXXXII-5;
	actXXXXXIII-1; actXXXXXIII-2; actXXXXXIII-3; actXXXXXIII-4; actXXXXXIII-5;
	actXXXXXIV-1; actXXXXXIV-2; actXXXXXIV-3; actXXXXXIV-4; actXXXXXIV-5;
	actXXXXXV-1; actXXXXXV-2; actXXXXXV-3; actXXXXXV-4; actXXXXXV-5;
	actXXXXXVI-1; actXXXXXVI-2; actXXXXXVI-3; actXXXXXVI-4; actXXXXXVI-5;
	actXXXXXVII-1; actXXXXXVII-2; actXXXXXVII-3; actXXXXXVII-4; actXXXXXVII-5;
	actXXXXXVIII-1; actXXXXXVIII-2; actXXXXXVIII-3; actXXXXXVIII-4; actXXXXXVIII-5;
	actXXXXXIX-1; actXXXXXIX-2; actXXXXXIX-3; actXXXXXIX-4; actXXXXXIX-5;
	actXXXXXX-1; actXXXXXX-2; actXXXXXX-3; actXXXXXX-4; actXXXXXX-5;
	actXXXXXXI-1; actXXXXXXI-2; actXXXXXXI-3; actXXXXXXI-4; actXXXXXXI-5;
	actXXXXXXII-1; actXXXXXXII-2; actXXXXXXII-3; actXXXXXXII-4; actXXXXXXII-5;
	actXXXXXXIII-1; actXXXXXXIII-2; actXXXXXXIII-3; actXXXXXXIII-4; actXXXXXXIII-5;
	actXXXXXXIV-1; actXXXXXXIV-2; actXXXXXXIV-3; actXXXXXXIV-4; actXXXXXXIV-5;
	actXXXXXXV-1; actXXXXXXV-2; actXXXXXXV-3; actXXXXXXV-4; actXXXXXXV-5;
	actXXXXXXVI-1; actXXXXXXVI-2; actXXXXXXVI-3; actXXXXXXVI-4; actXXXXXXVI-5;
	actXXXXXXVII-1; actXXXXXXVII-2; actXXXXXXVII-3; actXXXXXXVII-4; actXXXXXXVII-5;
	actXXXXXXVIII-1; actXXXXXXVIII-2; actXXXXXXVIII-3; actXXXXXXVIII-4; actXXXXXXVIII-5;
	actXXXXXXIX-1; actXXXXXXIX-2; actXXXXXXIX-3; actXXXXXXIX-4; actXXXXXXIX-5;
	actXXXXXXX-1; actXXXXXXX-2; actXXXXXXX-3; actXXXXXXX-4; actXXXXXXX-5;
	actXXXXXXXI-1; actXXXXXXXI-2; actXXXXXXXI-3; actXXXXXXXI-4; actXXXXXXXI-5;
	actXXXXXXXII-1; actXXXXXXXII-2; actXXXXXXXII-3; actXXXXXXXII-4; actXXXXXXXII-5;
	actXXXXXXXIII-1; actXXXXXXXIII-2; actXXXXXXXIII-3; actXXXXXXXIII-4; actXXXXXXXIII-5;
	actXXXXXXXIV-1; actXXXXXXXIV-2; actXXXXXXXIV-3; actXXXXXXXIV-4; actXXXXXXXIV-5;
	actXXXXXXXV-1; actXXXXXXXV-2; actXXXXXXXV-3; actXXXXXXXV-4; actXXXXXXXV-5;
	actXXXXXXXVI-1; actXXXXXXXVI-2; actXXXXXXXVI-3; actXXXXXXXVI-4; actXXXXXXXVI-5;
	actXXXXXXXVII-1; actXXXXXXXVII-2; actXXXXXXXVII-3; actXXXXXXXVII-4; actXXXXXXXVII-5;
	actXXXXXXXVIII-1; actXXXXXXXVIII-2; actXXXXXXXVIII-3; actXXXXXXXVIII-4; actXXXXXXXVIII-5;
	actXXXXXXXIX-1; actXXXXXXXIX-2; actXXXXXXXIX-3; actXXXXXXXIX-4; actXXXXXXXIX-5;
	actXXXXXXXI-1; actXXXXXXXI-2; actXXXXXXXI-3; actXXXXXXXI-4; actXXXXXXXI-5;
	actXXXXXXXII-1; actXXXXXXXII-2; actXXXXXXXII-3; actXXXXXXXII-4; actXXXXXXXII-5;
	actXXXXXXXIII-1; actXXXXXXXIII-2; actXXXXXXXIII-3; actXXXXXXXIII-4; actXXXXXXXIII-5;
	actXXXXXXXIV-1; actXXXXXXXIV-2; actXXXXXXXIV-3; actXXXXXXXIV-4; actXXXXXXXIV-5;
	actXXXXXXXV-1; actXXXXXXXV-2; actXXXXXXXV-3; actXXXXXXXV-4; actXXXXXXXV-5;
	actXXXXXXXVI-1; actXXXXXXXVI-2; actXXXXXXXVI-3; actXXXXXXXVI-4; actXXXXXXXVI-5;
	actXXXXXXXVII-1; actXXXXXXXVII-2; actXXXXXXXVII-3; actXXXXXXXVII-4; actXXXXXXXVII-5;
	actXXXXXXXVIII-1; actXXXXXXXVIII-2; actXXXXXXXVIII-3; actXXXXXXXVIII-4; actXXXXXXXVIII-5;
	actXXXXXXXIX-1; actXXXXXXXIX-2; actXXXXXXXIX-3; actXXXXXXXIX-4; actXXXXXXXIX-5;
	actXXXXXXXI-1; actXXXXXXXI-2; actXXXXXXXI-3; actXXXXXXXI-4; actXXXXXXXI-5;
	actXXXXXXXII-1; actXXXXXXXII-2; actXXXXXXXII-3; actXXXXXXXII-4; actXXXXXXXII-5

KEYWORDS
actII-1; actII-2; actII-3; actII-4; actIII; actIORF1; actIORF2; actIORF3; actIV; actVA1; actVA2; actVA3; actVA4; actVA5; actVA6; actVB; actVI; DNA-binding protein; gntK-family regulatory protein; integral membrane protein; lysR-family transcriptional regulatory protein; methyltransferase; mini-circle protein; mutr-like protein; oxidoreductase; regulatory protein; short-chain oxidoreductase.

SOURCE
Streptomyces coelicolor A3(2).
Streptomyces coelicolor A3(2).

REFERENCE
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 1 (bases 1 to 29080)
 Redenbach, M., Kieser, H. M., Hopwood, D. A.,
 Klenz, H., and Hopwood, D. A.
 A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome
 Mol. Microbiol. 21 (1), 77-96 (1996)

97000351 MEDLINE
8843436 PUBMED
2 (bases 1 to 29080)
REFERENCE
AUTHORS Warren, T. and Harris, D.

JOURNAL
REFERENCE
3 (bases 1 to 29080)
AUTHORS
Cerdeno, A.M., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
TITLE
Direct

JOURNAL Direct submission
Submitted (11-JUL-2001) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
On May 9, 2002 this sequence version replaced gi:14717081.
COMMENT

Structural variants coellicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics. Details of S. coellicolor sequencing at the Sanger Centre are available on the World Wide Web. (URL: <http://www.sanger.ac.uk/projects/S.coellicolor/>) CDS are numbered using the following system eg SC7B7.01c. SC (S. coellicolor), YB7 (cosmid name), .01 (first CDS), c (complementary

```

strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous. The length
in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon. Gene
prediction is based on positional base preference in codons using a
specially developed Hidden Markov Model (Krogh et al., Nucleic
Acids Research, 22(22):4768-4778(1994)) and the Frameplot program
of Bibb et al., Gene 30:157-66(1984) as implemented at
http://www.nih.go.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gtg, ttg or (att)) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions. Cosmid StBAC28G1
overlaps cosmids StBAC2Df6 and StBAC31E11.

FEATURES             source
    misc_feature      /organism="Streptomyces coelicolor A3(2)"
                     /strain="A3(2)"
                     /db_xref="taxon:100226"
                     /clone="cosmid BAC28G1"
    misc_feature      2..518
                     /note="previously sequenced DNA fragment. EMBL:X62373
                     S.coelicolor genes for hydroxycyl-CoA dehydrogenase"
    misc_feature      2..94
                     /note="nominal overlap with Streptomyces coelicolor cosmid
                     SCBAC2Df6"
    misc_feature      8..73
                     /note="PS01162 Quinone oxidoreductase / zeta-crystallin
                     signature"
    misc_feature      484..5883
                     /note="previously sequenced DNA fragment. EMBL:X58833.1
                     S.coelicolor 6 actVA region genes of the actinorhodin
                     biosynthetic gene cluster"
    gene              502..2103
                     /gene="SCBAC28G1.02"
                     /note="SCO5076; actVA1"
    CDS               502..2103
                     /gene="SCBAC28G1.02"
                     /note="SCBAC28G1.02, actVA1, probable integral membrane
                     protein, len: 533 aa; identical to previously sequenced
                     TR:Q53903 (EMBL:X58833) Streptomyces coelicolor 6 ActVA
                     region genes of the actinorhodin biosynthetic gene
                     cluster, ActVA1, 533 aa. Contains possible hydrophobic
                     membrane spanning regions"
                     /codon_start=1
                     /trans_table=11
                     /product="integral membrane protein"
                     /protein_id="CAC44189.1"
                     /db_xref="GI:I4717083"
                     /db_xref="SPTREMBL:Q53903"
                     /translation="MTANPRGGPGADQHPRWALGLVGLSLVLTNDVTNVLNTL
                     RTLDTPDGGLGASHSQVEYVLSAYTLAFATFLTWIGDRGLRRRLDLGLGLFGLS
                     SLAGAYASPEQLIARACMGVSGAAVLPLSTLATIAAVFLPRTGRPKALGTWIASVGFA
                     LIGPVFTGILLAHFWNGSVLPVLMAGLVAVLVPETRGTRAGRVDAAGLLLS
                     IAGVPLVYAIIEAGRSQGVTRPAVMAAGLAGLGILLVFLWHERRTPESLEIGFFPMR
                     KAFSTAAVAGVFYSFAMMGLFLSFAYFLVSQVGYTPIQAGCTTVALLAVANVCGPLST
                     VLVRISGPNNVCAAGMLAVNTASLCGYTFVTQHAPVNLILVLAALGAGVACVNPYTAIV
                     SINNAIPREKAGVASANNVTROLGALGVAVLGSMLGAARYRGIDEELAVLPSPSRH
                     QAGESLDATLLATRGESLGVPAQGFIDAMHLAAGAAAVALVGLAVLRWLPPSS
                     VTPTFPAGVAGREHSDDLKVOGS"
    RBS               2093..2097
                     /gene="SCBAC28G1.02"
    gene              2105..2500

```

Query Match	10.1%; Score 58.2; DB 1; Length 29080;
Best Local Similarity	49.5%; Pred. No. 10;
Matches 150; Conservative	0; Mismatches 153; Indels 0; Gaps 0;
QY	109 AGGAGGTGGACTACTTCTGCTCGGAGGAGGTGAGTTCGGTGGCCCGCGCGGTTCGGCGGCC 168
DB	17986 AGGAGGTCCGAAGCTGGCCGCGAGCTTCGCGGCGCGGTTCACGAGCGGCTCCGCCAGC 17927
QY	169 ACGCGGGGGGGCTTCAGCAGCAGCTGCTCAGGAGGAGTTCGAGAGGTTCGACACGCTCT 228
DB	17936 TCGTGCCTCGCTTACCTCGGGAGGCGCTGCAGACGGGGGACACGGCTGCTGCACGCTGCACCG 17867
QY	229 CACGCGCGGGGCCAAACACACACACCATGTGTACACGCGGGCGCACACGGCTTCGTGG 288
DB	17866 ACCTCGGGCACCCTCTCCCGCGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 17807
QY	289 TCGCGGAGCAGAGGTTCGAAGAGGACATCACACCTGCACCGCGGAGTTCACGAGCGCA 348
DB	17806 CCTCACCACCATGTGTGAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 17747
QY	349 GGGAGAGTTCCTCGCGAGGCGTAACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
DB	17746 AGCTGATGACGGTGCACCGGCGCGCTGACGGCTTCCTCGCGCGCGCGCGCGCGCGCG 17687
QY	409 GTG 411
DB	17686 GCG 17684
RESULT 10	
SCC30	18420 bp DNA linear BCT 12-MAY-2002
LOCUS	Streptomyces coelicolor cosmid C30.
DEFINITION	AL352972 AL645882
ACCESSION	AL352972.2 GI:20520831
VERSION	ABC transporter ATP-binding subunit; ABC transporter integral
KEYWORDS	membrane protein; arsc-family regulatory protein; bifunctional protein (ribonuclease H/phosphoglycerate mutase); dehydrogenase; integral membrane protein; KHG/KDPG aldolase; transmembrane transport protein; transport protein; two-component response regulator; two-component sensor kinase.
SOURCE	Streptomyces coelicolor A3(2).
ORGANISM	Streptomyces coelicolor A3(2).
REFERENCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
AUTHORS	Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces. 1 (bases 1 to 18420)
	Redenbach, M., Kieser, H.M., Denapaita, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.
	A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
TITLE	Mol. Microbiol. 21 (1), 77-96 (1996)
JOURNAL	97000351
MEDLINE	8843436
PUBLISHED	2 (bases 1 to 18420)
REFERENCE	Seeger, K.J. and Harris, D.
AUTHORS	Unpublished
JOURNAL	3 (bases 1 to 18420)
REFERENCE	Cerdeno, A.M., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
AUTHORS	Direct Submission
TITLE	Submitted (14-APR-2000) Streptomyces coelicolor sequencing project,
JOURNAL	

Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: bartrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK
On May 9, 2002 this sequence version replaced gi:7619747.

COMMENT

Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nh.gov.jp/jun/cgi-bin/frameplot.pl>.

CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid C30 Overlaps cosmid C75 on the AseI-C genomic restriction fragment.
Location/Qualifiers

FEATURES

source

1. 18420
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"

/clone="cosmid C30"

complement(346..771)

/gene="rnpB"

complement(346..771)

/gene="rnpB"

complement(346..771)

/note="probable ribonuclease P RNA, (rnpB) (positions taken from Streptomyces lividans EMBL:SLXLANB)"

complement(681..1661)

/gene="SCC30.01c"

/note="SCC02293"

complement(681..1661)

/gene="SCC30.01c"

/note="SCC30.01c, possible integral membrane protein, len: 326 aa; similar to TR:BAA87714 (EMBL:AB016260) Agrobacterium tumefaciens T1ORF89 protein, 306 aa; fasta scores: opt: 396 z-score: 465.4 E(): 1.8e-18; 28.7% identity in 286 aa overlap. Contains 2x Pfam match to entry PF00892 DUF6, integral membrane protein and possible hydrophobic membrane spanning regions"

/codon_start=1

/transl_table=11

/product="putative integral membrane protein"

/protein_id="CAB88171.1"

/db_xref="GI:7619748"

/db_xref="SPTREMBL:Q9L020"

/translation="MTARGWFLFSLMGVVGVYPMIKVAVDEVSPGVVFARCALGA ALLLPALRKGLLSTGTVRRRKMPLAFVIEIGPWWTLDARHLSSTAGLLIAGY PIVAVLLARFFGAERVGARRITGLGLAGVAVLTVPRLTGGDARSIAELVTVVGY ATAPLTAARHLKDVPTLQITPCLTLAAVYVAPAFRLPRLPSGEALAGLGLV CTALAFVALELIKVGTPRAGVIYVNPVAVAGALLDELTGLTGAVFTLLAG SVLTAAAGPGREARVPMSTRQTSRAGGRVESLTGLPFRPGSTGQDG"

misc_feature

complement(783..1190)

/gene="SCC30.01c"

/note="Pfam match to entry PF00892 DUF6, Integral membrane protein, score 51.40, E-value 1.9e-11"

complement(1212..1622)

/gene="SCC30.01c"

/note="Pfam match to entry PF00892 DUF6, Integral membrane protein, score 38.50, E-value 1.6e-07"

complement(1681..2517)

/gene="SCC30.02c"

/note="SCC02294"

complement(1681..2517)

/gene="SCC30.02c"

/note="SCC30.02c, probable araC-family regulatory protein, len: 278 aa; similar to SW:ARAC_SALTY (EMBL:J01797) Salmonella typhimurium arabinose operon regulatory protein AraC, 281 aa; fasta scores: opt: 232 z-score: 270.4 E(): 1.3e-07; 26.7% identity in 270 aa overlap and to various Streptomyces coelicolor putative araC-family transcriptional regulators, e.g. TR:O50480 (EMBL:AL031031) S. coelicolor putative transcriptional regulator SC7C7.17 (SC4H8.01), 288 aa; fasta scores: opt: 638 z-score: 605.1 E(): 3.3e-28; 43.8% identity in 292 aa overlap. Contains Pfam match to entry PF00165 HTH_AraC, Bacterial regulatory helix-turn-helix proteins, araC family and possible helix-turn-helix motif at residues 189..210 (+3.81 SD)"

/codon_start=1

/transl_table=11

/product="putative araC-family regulatory protein"

/protein_id="CAB88172.1"

/db_xref="GI:7619749"

/db_xref="SPTREMBL:Q9L019"

/translation="MAGAREQVGRARLRGRGPPPLDLTLTARFDRHYVAPHARDEVTV GVTGVEIVAYRGHISGPGSIVVLEPGVHTGPGAPGEGYSYALYAAARLLTGD TKTDTVLHFRFVLDPPDLAALRAAHTDLARCPLAESRLPMLTALARRGSTA RAADTVPGAGRVAVRDRDLAPSLAALADLGLSLRYOLLRFRTSTGHPY AWLAQHRVARARGLLDAGLRPAEVALVGFADQALTRWFRVLTGVTATYRNSVQDR TG"

complement(1699..1956)

/gene="SCC30.02c"

/note="Pfam match to entry PF00165 HTH_AraC, Bacterial regulatory helix-turn-helix proteins, araC family, score 59.80, E-value 5.8e-14"

2628..3389

/gene="SCC30.03"

/note="SCC02295"

2628..3389

/gene="SCC30.03"

/note="SCC30.03"

/note="SCC30.03, hypothetical protein, len: 253 aa; N-terminal region similar to TR:CAB52842 (EMBL:AL109848) Streptomyces coelicolor putative transcriptional regulator SC151.18, 548 aa; fasta scores: opt: 114 z-score: 134.1 E(): 5; 30.6% identity in 196 aa overlap. Contains possible coiled-coil region at approx residues 76..93"

/codon_start=1

/transl_table=11

/product="hypothetical protein SCC30.03"

/protein_id="CAB88173.1"

/db_xref="GI:7619750"

/db_xref="SPTREMBL:Q9L018"

/translation="MRIGELADTVGTVTRTVRHHQGLLPEPERLANGRYRTVLRHV VVLARIRRLTGLGLAEVRDVLADDAKDLAEVLELSDLSLARQEAIRERRTLRA LLGTGGVPGEGVSPLEALFAGIDVSPNALRDREMLVLESLTVAPEERAGLLA AMRGALGSPAARARRTYELLDELADAGADDPVAEAVRLADCPAELLPEGFDL DPGHLLRLYADFAPAQAQAEVRAEMEIAARGRS"

3386..4000

/gene="SCC30.04"

/note="SCC02296"

3386..4000

/gene="SCC30.04"

/note="SCC30.04, possible integral membrane protein, len: 204 aa; similar to TR:AAF12504 (EMBL:AE001863) Deinococcus radiodurans hypothetical 32.0 kD protein, 289 aa; fasta

RESULT 13					
AP004045/c					
LOCUS			DNA	linear	HTG 21-MAR-2002
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 2 clone OJ1135_F06, *** SEQUENCING IN PROGRESS ***, in ordered pieces.				
ACCESSION	AF004045				
VERSION	AF004045.1	GI:15208413			
KEYWORDS	HTG; HTGS_PHASE2.				
SOURCE	Oryza sativa (japonica cultivar-group)	(cultivar:Nipponbare)	DNA,		
	clone:OJ1135_F06.				
ORGANISM	Oryza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Ehrhartoideae; Oryzeae; Oryza.				
REFERENCE					
	1				

Search completed: June 16, 2003, 03:10:51
Job time : 1694.35 secs

AUTHORS Rafanan,E.R. Jr., Le,L., Zhao,L., Decker,H. and Shen,B.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAY-2000) Chemistry, University of California - Davis, One Shields Ave., Davis, CA 95616, USA

FEATURES
 source 1. 3164
 /organism="Streptomyces olivaceus"
 /strain="Tu 2353"
 /db_xref="taxon:47716"
 339. .1958
 /gene="elmC"
 339. .1958
 /gene="elmC"
 /note="ElmC"
 /codon_start=1
 /transl_table=11
 /product="Tcm B2 oxygenase"
 /protein_id="AAF73050.1"
 /db_xref="GI:8118601"
 /translation="MDRIEIPVLVGGGLTGLAAAVFLRQGVDCLLVERHRRSTTFLT
 RASIVTEPDLVADVPSTQAWCGQKLEPILRDEAVRGGADIRFTRLDLSFAQAD
 GVDIVDRGFTYANQAPGLVTFDENWIFGFCDAYGGGAAFTGRCADIVRT
 GLGPDIPDLQVLPQMSHHVADSVSRGTVFLAGDAAHVHPAGAFGANGGIQDAH
 NLAKLASLVHGRASDALLDTYHQERHPVGTETAEQAQWTRHYRLDGGDELGRILDT
 KVAAGYRTSSAVLGAAYPATIPHEIALTLPGORPHWLDHDDGGRKVSIVDLAVDG
 FVLLARADGTPWADAAARLAATTGIPLTAHVYKSTLTDPADALAAATGLGEGAGALLR
 PDGFVARSDDTSADDPPEAVLDGVLARILART"
 1990. .2319
 /gene="elmH"
 1990. .2319
 /gene="elmH"
 /note="ElmH"
 /codon_start=1
 /transl_table=11
 /product="Tcm F1 monooxygenase"
 /protein_id="AAF73051.1"
 /db_xref="GI:8118602"
 /translation="MPLRSPDRAHYTLNVFTVSPDDOEKLYRLHVDVTEIRHLPG
 FYSATFHNGNRHVNVYAQWSEETFRSMHALPELQEHFAFCRGCIATPLSVPCAVAL
 VFEEAAG"
 2316. .2636
 /gene="elmi"
 2316. .2636
 /gene="elmi"
 /note="Elmi"
 /codon_start=1
 /transl_table=11
 /product="Tcm F2 cyclase"
 /protein_id="AAF73052.1"
 /db_xref="GI:8118603"
 /translation="MTYRALMVLMDPADAERKVAAFEDHDTTGLPEEIGLRRRTLFR
 FHDLYMHLEADDDVLPLKLEARDPPFRKVNVEEVRGLYTPVASDWSRLTDSKAEVIV
 TWEA"
 2648. .3079
 /gene="elmJ"
 2648. .3079
 /gene="elmJ"
 /note="may effect polyketide antibiotic elloramycin A
 biosynthesis"
 /codon_start=1
 /transl_table=11
 /product="ElmJ"
 /protein_id="AAF73053.1"
 /db_xref="GI:8118604"
 /translation="MTAQQPPPIGVDDVTPSTQGRTRALLTPTSAGATGGTGLTL
 DLGGERISHVHPYDKILYAVGVSVEVDGHEITLGTDOALLTRGMRHRRFHNRT
 EARFLYQISPLAPRDLGLHVDTEVPREPQDAPPTVGGVR"
 487 a 1236 c 1003 g 437 t 1 others

BASE COUNT
 ORIGIN
 Query Match 9.5%; Score 54.6; DB 1; Length 3164;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 05:14:00 ; Search time 1163.61 Seconds
(without alignments)
7989.092 Million cell updates/sec

Title: US-10-090-035-3
Perfect score: 574
Sequence: 1 accacgcgtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 574

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.*

1: em_estba.*
2: em_estchum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	417	72.6	417	13	BM501439
2	390	67.9	524	10	AW288876
3	387	67.4	500	10	AW787732
4	363	63.2	465	10	BE129897
5	340	59.2	523	10	AW331212
6	315	54.9	552	10	BE025302

7	283	49.3	436	10	AW787315
8	254	44.3	326	10	BE025303
9	251	43.7	443	10	BE129644
10	223	38.9	274	10	AW288875
11	210	36.6	476	10	BE519299
12	207	36.1	311	12	BF729420
13	200	34.8	496	10	AW787314
14	178	31.0	648	14	BQ619167
15	178	31.0	648	14	BQ619315
16	178	31.0	648	14	BQ619318
17	178	31.0	648	14	BQ619337
18	178	31.0	648	14	BQ619383
19	178	31.0	648	14	BQ619390
20	174	30.3	553	11	AY104409
21	167	29.1	433	10	BE225008
22	157	27.4	226	9	AI372183
23	154	26.8	550	12	BG840383
24	153	26.7	546	9	AA979839
25	143	24.9	401	10	AW289056
26	139	24.2	466	9	AI964534
27	139	24.2	481	9	AI964458
28	139	24.2	539	9	AI855425
29	85	14.8	168	14	T12683
30	85	14.8	201	9	AA051890
31	85	14.8	215	14	T18841
32	85	14.8	230	9	AA051874
33	85	14.8	240	14	W21677
34	52	9.1	188	10	AW679347
35	52	9.1	287	10	BE125310
36	52	9.1	288	10	BE125500
37	52	9.1	348	10	BE593507
38	52	9.1	399	13	BM318672
39	52	9.1	466	10	BE597738
40	52	9.1	485	10	BE599123
41	52	9.1	509	14	BQ280894
42	52	9.1	513	14	BQ280709
43	52	9.1	516	10	BQ364814
44	52	9.1	519	10	AW680016
45	52	9.1	522	10	AW679969

ALIGNMENTS

RESULT 1
BM501439
LOCUS BM501439 417 bp mRNA linear EST 14-FEB-2002
DEFINITION PAC000000000593 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
ACCESSION BM501439
VERSION BM501439.1 GI:18661517
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 417)
Hunter,B.G., Beatty,M., Singletary,G., Hamaker,B., Larkins,B.A. and Jung,R.

AUTHORS

Maize opaque endosperm mutations create extensive changes in patterns of gene expression
Unpublished (2002)
Contact: Jung R

TITLE

Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619

JOURNAL

COMMENT

FEATURES

source

/db_xref="taxon:4577"

BASE COUNT		81 a 150 c 133 g 53 t		/clone_lib="pioneer AF-1 array"	
ORIGIN				/note="Vector: pSport1; site_1: SalI; site_2: NotI"	
Query Match		72.6%; Score 417; DB 13; Length 417;			
Best Local Similarity		100.0%; Pred. No. 0;			
Matches 417; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
QY	12	CGCCACGGCTCGGCACAGCAATCCACACAAGCACTTCGACGTCACACGGCGGCTCGCGCA	71		
DB	1	CGCCACGGCTCGGCACAGCAATCCACACAAGCACTTCGACGTCACACGGCGGCTCGCGCA	60		
QY	72	CAGACACACAACGCTCGGCACCAATGGCTTACTACGAGAGGTGGACTACTGCTCGGAG	131		
DB	61	CAGACACACAACGCTCGGCACCAATGGCTTACTACGAGAGGTGGACTACTGCTCGGAG	120		
QY	132	GAGGTGAGTCGTGGCCGGCGGCTTCGGCGGCACGGCGGCGGCTCCAGCAGCAC	191		
DB	121	GAGGTGAGTCGTGGCCGGCGGCTTCGGCGGCACGGCGGCGGCTCCAGCAGCAC	180		
QY	192	GTGCTCAAGGAGAGTTCGAGAGGTGCACACGGTCTCAGCCGCCGGCGCAACCCAC	251		
DB	181	GTGCTCAAGGAGAGTTCGAGAGGTGCACACGGTCTCAGCCGCCGGCGCAACCCAC	240		
QY	252	CACCACCATGTCACACGCGCGGCTTCGTGTGCGCAGAGCAGAGGTCGGAAGAG	311		
DB	241	CACCACCATGTCACACGCGCGGCTTCGTGTGCGCAGAGCAGAGGTCGGAAGAG	300		
QY	312	GACATCAACACCTCACCAGCGAGGTCACGAGCGCAGGAGAGCTTCCTCGCCAGGGCT	371		
DB	301	GACATCAACACCTCACCAGCGAGGTCACGAGCGCAGGAGAGCTTCCTCGCCAGGGCT	360		
QY	372	AACGTAGCGCCCGCGCGGATCCACGCCGTTCTGCTTCCTCGCTGCGCTTA	428		
DB	361	AACGTAGCGCCCGCGCGGATCCACGCCGTTCTGCTTCCTCGCTGCGCTTA	417		
RESULT 2					
AW288876					
LOCUS 707009507.x4 707 - Mixed adult tissues from Walbot lab (SK) Zea					
DEFINITION 524 bp mRNA linear EST 16-JAN-2000					
mays cDNA, mRNA sequence.					
ACCESSION AW288876					
VERSION AW288876.1 GI:6695663					
KEYWORDS EST.					
SOURCE Zea mays.					
ORGANISM Zea mays					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC					
clade; Panicoideae; Andropogoneae; Zea.					
1. (bases 1 to 524)					
Walbot V.					
Maize ESTs from various cDNA libraries sequenced at Stanford					
University					
Unpublished (1999)					
Contact: Walbot V					
Department of Biological Sciences					
Stanford University					
855 California Ave, Palo Alto, CA 94304, USA					
Tel: 650 723 2227					
Fax: 650 725 8221					
Email: walbot@stanford.edu					
Plate: 707009 row: E column: 07.					
Location/Qualifiers					
1. .524					
/organism="Zea mays"					
/cultivar="W23"					
/db_xref="taxon:4577"					
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"					
/tissue_type="tassel, kernel, silk, husk, root, leaf"					
/dev stage="adult"					
)					
FEATURES					
source					

Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. .500

FEATURES

source

/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 97 a 159 c 156 g 88 t

Query Match 67.4%; Score 387; DB 10; Length 500;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 62 CGCTGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTGACTA 121
DB 1 CGCTGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTGACTA 60
QY 122 CTGCTCGGAGGAGTGAGTGGTGGCCCGCGCGGCTTCGGCCGCCACGCGCGCGCGT 181
DB 61 CTGCTCGGAGGAGTGAGTGGTGGCCCGCGCGGCTTCGGCCGCCACGCGCGCGCGT 120
QY 182 CCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCGCGC 241
DB 121 CCAGCAGCAGCTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCGCGC 180
QY 242 CAACACACACACACACCATGTCACACCGCGCGCGCGCTTCGTGTGTCGCGAGACGAG 301
DB 181 CAACACACACACACACCATGTCACACCGCGCGCGCGCTTCGTGTGTCGCGAGACGAG 240
QY 302 GGTGAGAGGACATCAACACCTGACCGCGCGGAGTCCAGAGCGAGGAGAGCTTCCT 361
DB 241 GGTGAGAGGACATCAACACCTGACCGCGCGGAGTTCAGAGCGAGGAGAGCTTCCT 300
QY 362 CGCCAGGCGTAACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 421
DB 301 CGCCAGGCGTAACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 422 TGCCTTATGATGCTGTGTGTGACTGGTTGTTTCAGGGTTCATCGTACTTGGCTATCGTAC 481
DB 361 TGCCTTATGATGCTGTGTGTGACTGGTTGTTTCAGGGTTCATCGTACTTGGCTATCGTAC 420
QY 482 GTGCACGACCTAGCTCTCTAGCAATACGACAATAAGCTGTCACCTGAATAAACTT 541
DB 421 GTGCACGACCTAGCTCTCTAGCAATACGACAATAAGCTGTCACCTGAATAAACTT 480
QY 542 CTTGCGTAAT 550
DB 481 CTTGCGTAAT 489

RESULT 4

BE129897

LOCUS

DEFINITION BE129897 945032C12.X1 945 - Mixed adult tissues from Walbot lab, same as 707 (SK) Zea mays cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 465)
Walbot, V.

Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 945032 row: C column: 12.

FEATURES

source

Location/Qualifiers
1. .465

/organism="Zea mays"
/cultivar="W23"

/db_xref="taxon:4577"
/clone_lib="945" - Mixed adult tissues from Walbot lab,
same as 707 (SK)
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"

/note="Organ: tassel, kernal, silk, husk, root, leaf;
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT 88 a 151 c 149 g 77 t

BASE COUNT

Query Match 63.2%; Score 363; DB 10; Length 465;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 57 CACGGCGCTGGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTG 116
DB 1 CACGGCGCTGGCGCAGACACACCAAGCGTCGGCACCACCAATGGCTTACTACAGGAGGTG 60
QY 117 GACTACTGCTCGGAGGAGTGAGTGGTGGCCCGCGCGCGCGCTTCGGCGCGCACGCGCGC 176
DB 61 GACTACTGCTCGGAGGAGTGAGTGGTGGCCCGCGCGCGCTTCGGCGCGCACGCGCGC 120
QY 177 GCGGTCACGACGACGTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCC 236
DB 121 GCGGTCACGACGACGTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCC 180
QY 237 GCGGTCACGACGACGTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCC 296
DB 181 GCGGTCACGACGACGTCGTCAAGGAGAAGTTCGAGGAGGTGCGACACGGTCTCACGCGCC 240
QY 297 ACCAGGGTCGAGAGGAGCATCAACACCTGACCGCGCGAGGTCCAGAGCGGAGGAGAGC 356
DB 241 ACCAGGGTCGAGAGGAGCATCAACACCTGACCGCGCGAGGTCCAGAGCGGAGGAGAGC 300
QY 357 TTCCTCGCCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCATCCAGCGCGGTTCGTCGCTTCG 416
DB 301 TTCCTCGCCAGGGCTAACTGAGCGCGCGCGCGCGCGCGCATCCAGCGCGGTTCGTCGCTTCG 360
QY 417 CTGGTGCTTATGATGCTGTGTGTGACTGGTTGTTTCAGGGTTCATCGTACTTGGCTAT 476
DB 361 CTGGTGCTTATGATGCTGTGTGTGACTGGTTGTTTCAGGGTTCATCGTACTTGGCTAT 420
QY 477 CGTACGTGACGCACTCAGCTCCTGTACGAAATACGACAATAAGC 521
DB 421 CGTACGTGACGCACTCAGCTCCTGTACGAAATACGACAATAAGC 465

RESULT 5

Mon Jun 16 14:55:56 2003

```

AW331212      523 bp  mRNA  linear  EST 31-JAN-2000
LOCUS         707049B04.x1 707 - Mixed adult tissues from Walbot lab (SK) Zea
DEFINITION    mays cDNA, mRNA sequence.
ACCESSION     AW331212
VERSION       AW331212.1 GI:6827569
KEYWORDS      EST.
SOURCE        Zea mays.
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 523)
REFERENCE     Walbot,V.
AUTHORS       Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE         University
              Unpublished (1999)
JOURNAL       Contact: Walbot V
COMMENT       Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 707049 row: E column: 04.
              Location/Qualifiers
              1..523
              /organism="Zea mays"
              /cultivar="W23"
              /db_xref="taxon:4577"
              /clone_lib="707 - Mixed adult tissues from Walbot lab (SK
              )"
              /tissue_type="tassel, kernel, silk, husk, root, leaf"
              /dev_stage="adult"
              /lab_host="DHI08"
              /note="Organ: tassel, kernel, silk, husk, root, leaf;
              Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
              differentiated maize tissues from an active Mutator
              plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
              husk, root, leaf). Unidirectionally cloned."
BASE COUNT    125 a 158 c 149 g 91 t
ORIGIN
Query Match      59.2%; Score 340; DB 10; Length 523;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 490; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 63 CGTGGCGCAGACACACCAAGCGTGGGACCAATGGCTTACTACAGGAGTGGACTAC 122
DB 2 CGCTGCGCAGACACACCAAGCGTGGGACCAATGGCTTACTACAGGAGTGGACTAC 61
QY 123 TGCTCGGAGGAGTGGCTGGGCGCGCGCGGCTTCGGCGCCACGCGCGCGCGTC 182
DB 62 TGCTCGGAGGAGTGGCTGGGCGCGCGCGGCTTCGGCGCCACGCGCGCGGTC 121
QY 183 CAGCAGCAGCTGCTCAAGAGAGAGTTCGAGAGAGTCACACGGTCTCACCGCGCGCGGCC 242
DB 122 CAGCAGCAGCTGCTCAAGAGAGAGTTCGAGAGAGTCACACGGTCTCACCGCGCGGCC 181
QY 243 AACCACACACACACACCTGTCACACGCGCGCGCGGCTTCGTTGGTCCGAGACAGG 302
DB 182 AACCACACACACACACCTGTCACACGCGCGCGCGGCTTCGTTGGTCCGAGACAGG 241
QY 303 GTCGAAGAGGACATCAACACCTGTCACCGCGCGAGTCCACGCGCAGGAGAGTTCCTC 362
DB 242 GTCGAAGAGGACATCAACACCTGTCACCGCGCGAGTTCACGCGCAGGAGAGTTCCTC 301
QY 363 GCCAGGCGTACTGAGCGCGCGCGCGCGCATCCAGCGCGTTCGTTGCTGCTCGCT 422
DB 302 GCCAGGCGTACTGAGCGCGCGCGCGCGCATCCAGCGCGTTCGTTGCTGCTCGCT 361
QY 423 GCCTTATGTATGTCTGTGGTGTGACTGTGTTGTTGAGGCTCATCGTACTTGGCTATCGTACG 482

```

```

362 GCCTTATGTATGTCTGTGGTGTGACTGTGTCAGGGGTCACTGCTTGGCTATCGTACG 421
483 TGCAGGCACCTCAGCTCTCTGTACGAATACACAAATAGCTCGTACCTGAATAAACTTC 542
DB 422 TGCAGGCACCTCAGCTCTCTGTACGAATACACAAATAGCTCGTACCTGAATAAACTTC 481
QY 543 TTCGTAATACTAA 555
DB 482 TTCGTAATACTAA 494

RESULT 6
BE025302      552 bp  mRNA  linear  EST 07-JUN-2000
LOCUS         945028B09.Y1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION    (SK) Zea mays cDNA, mRNA sequence.
ACCESSION     BE025302
VERSION       BE025302.1 GI:8318737
KEYWORDS      EST.
SOURCE        Zea mays.
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 552)
REFERENCE     Walbot,V.
AUTHORS       Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE         University
              Unpublished (1999)
JOURNAL       Contact: Walbot V
COMMENT       Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 945028 row: B column: 09.
              Location/Qualifiers
              1..552
              /organism="Zea mays"
              /cultivar="W23"
              /db_xref="taxon:4577"
              /clone_lib="945 - Mixed adult tissues from Walbot lab,
              same as 707 (SK)"
              /tissue_type="tassel, kernel, silk, husk, root, leaf"
              /dev_stage="fully-grown"
              /lab_host="DHI08"
              /note="Organ: tassel, kernel, silk, husk, root, leaf;
              Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
              differentiated maize tissues from an active Mutator plant.
              Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
              root, leaf). Unidirectionally cloned. New library number
              given to library 707 for additional sequencing."
BASE COUNT    135 a 165 c 160 g 92 t
ORIGIN
Query Match      54.9%; Score 315; DB 10; Length 552;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 515; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 49 CGACGTACACGGCGCTCGGCACACACACACACAGCGTCGGCACCATTGGCTTACTAC 108
DB 1 CGACGTACACGGCGCTCGGCACACACACACACAGCGTCGGCACCATTGGCTTACTAC 60
QY 109 AGGAGTGGAGTGGCTCGGAGAGGTGAGTGGCGCGCGCGCGCTTCGGCGGCC 168
DB 61 AGGAGTGGAGTGGCTCGGAGAGGTGAGTGGCGCGCGCGCTTCGGCGGCC 120
QY 169 ACCGCGCGCGCTCCAGCAGCAGCTCGTCAAGAGAGAGTTCGAGAGGTTCGACACGGTCT 228
DB 121 ACCGCGCGCGCTCCAGCAGCAGCTCGTCAAGAGAGAGTTCGAGAGGTTCATACGGTCT 180
QY 229 CACGCGCGCGCGCCACACACCACCATGGTTCACCACGCGCGCGCGCTTCGTGG 288

```

```

181 CACGCGCGGCGCCAAACACACACACACACATGGTCACCGCGCCACACGGCTTCGTGG 240
182 |||||||
289 TCGCGGAGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGGAGGTCACAGCGCA 348
290 |||||||
241 TCGCGGAGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGGAGGTCACAGCGCA 300
242 |||||||
349 GGGAGAGCTTCTCGCCAGGCTAACTGAGCGCGCCGCGCGGCATCCACGCGGTC 408
350 |||||||
301 GGGAGAGCTTCTCGCCAGGCTAACTGAGCGCGCGCCGCGGCATCCACGCGGTC 360
302 |||||||
409 GTGCTTGCTCGGTCCTTATGTATGTCTGTGGTGTACTGGTTGTCAGGTCATCGTAC 468
410 |||||||
361 GTGCTTGCTCGGTCCTTATGTATGTCTGTGGTGTACTGGTTGTCAGGTCATCGTAC 420
362 |||||||
469 TTGGCTATCTGACGTGACGACACCTGAGCTCTGTGACGATTAACGACATTAAGTCGTGAC 528
470 |||||||
421 TTGGCTATCTGACGTGACGACACCTGAGCTCTGTGACGATTAACGACATTAAGTCGTGAC 480
422 |||||||
529 CTGAATAAACTTCTTCGTAATACTAAATAAAAAAAAAAAAAA 567
530 |||||||
481 CTGAATAAACTTCTTCGTATATACTAAATAAAAAAAAAAAAAA 519
482 |||||||

```

```

RESULT 7
AW787315
LOCUS
DEFINITION
436 bp mRNA linear EST 16-MAY-2000
(SK) Zea mays cDNA, mRNA sequence.
ACCESSION
AW787315
VERSION
AW787315.1 GI:7844112
KEYWORDS
EST.
SOURCE
ze mays.
ORGANISM
ze mays

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945002 row: E column: 06.
Location/Qualifiers
1. 436
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"

```

```

note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT
82 a 136 c 136 g 82 t
ORIGIN
Query Match 49.3%; Score 283; DB 10; Length 436;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 433; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY
115 TGGACTACTGCTCGAGAGAGGTGAGGTGCGTGGCCCGCGCGGCTTCGGCGCCACGCGC 174
116 |||||||
Db
1 TGGACTACTGCTCGAGAGAGGTGAGGTGCGTGGCCCGCGCGGCTTCGGCGCCACGCGC 60
117 |||||||
OY
175 GCGCGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTTCGAGGAGGTTCGACACGGTCTTCACGCG 234
176 |||||||
Db
61 GCGCGCTCCAGCAGCAGCTCGTCAAGGAGAAGTTTCGAGGAGGTTCGACACGGTCTTCACGCG 120
62 |||||||
OY
235 CCGCGCGCAACCCACACACACCATGCTACACAGCGCGGCGGCTTCGTTGTGTGCGCGC 294
236 |||||||
Db
121 CCGCGCGCAACCATCACCACCATGCTACACAGCGCGGCGGCTTCGTTGTGTGCGCGC 180
122 |||||||
OY
295 AGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGGAGGTCCACGAGCGCAGGAGA 354
296 |||||||
Db
181 AGACACAGGTCGAAGAGGACATCAACACCTGCACCGCGGAGGTCCACGAGCGCAGGAGA 240
182 |||||||
OY
355 GCTTCTCTCGCGAGGCTAACTGAGCGCGCGCGCGGCGGCGGATCCACCGCGGCTTCGTCGCTT 414
356 |||||||
Db
241 GCTTCTCTCGCGAGGCTAACTGAGCGCGCGCGCGGCGGCGGATCCACCGCGGCTTCGTCGCTT 300
242 |||||||
OY
415 GCCTGCGTGCCTTATGTATGTCTGTGGTGTACTGGTTGTCAGGTCATCGTACTTGGCT 474
416 |||||||
Db
301 GCCTGCGTGCCTTATGTATGTCTGTGGTGTACTGGTTGTCAGGTCATCGTACTTGGCT 360
302 |||||||
OY
475 ATCGTAGTGTGACGACCTGAGCTCTGTGACGAATTACGACATTAAGTCGTCGACCTCAAT 534
476 |||||||
Db
361 ATCGTAGTGTGACGACCTGAGCTCTGTGACGAATTACGACATTAAGTCGTCGACCTCAAT 420
362 |||||||
OY
535 AAAACTTCTTCGTAAT 550
536 |||||||
Db
421 AAAACTTCTTCGTAAT 436
422 |||||||

```

```

RESULT 8
BE025303
LOCUS
DEFINITION
326 bp mRNA linear EST 07-JUN-2000
945028B09.y2 945 - Mixed adult tissues from Walbot lab, same as 707
(SK) Zea mays cDNA, mRNA sequence.
ACCESSION
BE025303
VERSION
BE025303.1 GI:8318738
KEYWORDS
EST.
SOURCE
ze mays.
ORGANISM
ze mays

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945028 row: B column: 09.
Location/Qualifiers
1. 326
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"

```

```

note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT
82 a 136 c 136 g 82 t
ORIGIN
Query Match 49.3%; Score 283; DB 10; Length 436;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 433; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

FEATURES
source
1. 326
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"

```

```

note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."
BASE COUNT
82 a 136 c 136 g 82 t
ORIGIN
Query Match 49.3%; Score 283; DB 10; Length 436;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 433; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```


differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned.

BASE COUNT 57 a 91 c 93 g 33 t
ORIGIN

Query Match 38.9%; Score 223; DB 10; Length 274;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CAAGCACATTCAGCTCACACGGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 99
Db 1 CAAGCACATTCAGCTCACACGGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 60
QY 100 CTTACTACAGGAGGTGACTCTCTCGAGGAGGTGAGTTCGTCGGCCCGCGCGCT 159
Db 61 CTTACTACAGGAGGTGACTCTCTCGAGGAGGTGAGTTCGTCGGCCCGCGCGCT 120
QY 160 TCGGCGCCACAGGCGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 219
Db 121 TCGGCGCCACAGGCGCGCTCGGCACACACACACAAAGCTCGGCACCAATGG 180
QY 220 ACACGGCTCTACGCGCGCGCTCGGCACACACACACCAATGG 262
Db 181 ACACGGCTCTACGCGCGCGCTCGGCACACACACCAATGG 223

RESULT 11
LOCUS BE519299
DEFINITION 945007B05.x1 945 - Mixed adult tissues from Walbot lab, same as 707
ACCESSION BE519299
VERSION BE519299.1 GI:9743151
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 476)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945007 row: B column: 05.
Location/Qualifiers
1. 476
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; vector: pGAD10; Site: 1: EORI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."

BASE COUNT 99 a 145 c 146 g 86 t
ORIGIN

Query Match 36.6%; Score 210; DB 10; Length 476;

Best Local Similarity 98.9%; Pred. No. 0;
Matches 460; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 98 GGCTTACTTACCAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTTCGGTGGCCCGCGCG 157
Db 12 GGCTTACTTACCAGGAGGTGGACTACTGCTCGAGGAGGTGAGGTTCGGTGGCCCGCGCG 71
QY 158 CTTGCGCGCCACCGCGCGCGCTCGGCACACACACCAATGGTTCGAGGAGGT 217
Db 72 CTTGCGCGCCACCGCGCGCGCTCGGCACACACACCAATGGTTCGAGGAGGT 131
QY 218 CGACACGGTCTCACCGCGCGCGCTCGGCACACACCAATGGTTCGAGGAGGT 277
Db 132 CGACACGGTCTCACCGCGCGCGCTCGGCACACCAATGGTTCGAGGAGGT 191
QY 278 CGGCTTCTGCTGCGCGCGCGCTCGGCACACCAATGGTTCGAGGAGGT 337
Db 192 CGGCTTCTGCTGCGCGCGCGCTCGGCACACCAATGGTTCGAGGAGGT 251
QY 338 CCAGCAGCGGAGGAGGTTCCTCGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 397
Db 252 CCAGCAGCGGAGGAGGTTCCTCGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 311
QY 398 CACGCGGTTCTGCTGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 457
Db 312 CACGCGGTTCTGCTGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 371
QY 458 GGTCTATGCTTGGCTATCTGCTGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 517
Db 372 GGTCTATGCTTGGCTATCTGCTGCGCGCGCTCGGCACCAATGGTTCGAGGAGGT 431
QY 518 AAGCTGTGACCTGTAATAAATCTCTGTAATAAATCTCTGTAATAAATCTCTGTAATAA 562
Db 432 AAGCTGTGACCTGTAATAAATCTCTGTAATAAATCTCTGTAATAAATCTCTGTAATAA 476

RESULT 12
LOCUS BF729420
DEFINITION 1000077C04.x2 1000 - Unigene I from Maize Genome Project Zea mays
ACCESSION BF729420
VERSION BF729420.1 GI:12047281
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 311)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1000077 row: C column: 04.
Location/Qualifiers
1. 311
/organism="Zea mays"
/db_xref="dbEST:945032C12.Y1"
/db_xref="taxon:4577"
/clone_lib="1000 - Unigene I from Maize Genome Project"
/note="This library represents the unique ESTs found in the first round of EST sequencing at Stanford University for the maize genome project. Sequences are present from libraries 486, 487, 496, 603, 606, 614, 618, 660, 683, 687, 707, and 945. Contigs were assembled using TIGR's

FEATURES
source

CAP program and a representative EST from each contig was selected for the Unigene set. All singlets were also selected."

```

BASE COUNT      53 a  105 c  107 g  46 t
ORIGIN

Query Match      36.1%; Score 207; DB 12; Length 311;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 183 CAGCAGCAGCTGTCGAAGGAGAAGTTCGAGGAGTTCGACACGCTCTCAGCGCCGCCGCC 242
      |||||||
DB 54 CAGCAGCAGCTGTCGAAGGAGAAGTTCGAGGAGTTCGACACGCTCTCAGCGCCGCCGCC 113
      |||||||
QY 243 AACCCACACACCATGTCACACGCGCGCGGCTTCGTCGCGCGAGACGAG 302
      |||||||
DB 114 AACCCACACACCATGTCACACGCGCGCGGCTTCGTCGCGCGAGACGAG 173
      |||||||
QY 303 GTCGAAGAGACATCAACACCTGCACCGCGGAGTCCACGAGCGCAGGAGAGCTTCCTC 362
      |||||||
DB 174 GTCGAAGAGACATCAACACCTGCACCGCGGAGTTCACGAGCGCAGGAGAGCTTCCTC 233
      |||||||
QY 363 GCCAGGCTAACTAGCGCGCGCGCGCGCGGCTTCACGCGCGCTTCGCTGCTGCGCTGCCT 422
      |||||||
DB 234 GCCAGGCTAACTAGCGCGCGCGCGCGCGGCTTCACGCGCGCTTCGCTGCTGCGCTGCCT 293
      |||||||
QY 423 GCCTATGATGTCGTG 440
      |||||||
DB 294 GCCTATGATGTCGTG 311
      |||||||

```

```

RESULT 13
AW787314.
LOCUS      496 bp mRNA linear EST 16-MAY-2000
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION AW787314
VERSION    1
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
REFERENCE  1 (bases 1 to 496)
AUTHORS    Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 945002 row: E column: 06.
FEATURES   source
            1..496
            Location/Qualifiers
            1..496
            /organism="Zea mays"
            /cultivar="W23"
            /db_xref="taxon:4577"
            /clone_lib="945 - Mixed adult tissues from Walbot lab,
            same as 707 (SK)"
            /tissue_type="tassel, kernal, silk, husk, root, leaf"
            /dev_stage="fully-grown"
            /lab_host="DH10B"
            /note="Organ: tassel, kernal, silk, husk, root, leaf;
            Vector: pGAD10; Site: EcoRI; cDNA library from fully
            differentiated maize tissues from an active Mutator plant.
            Tissue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
            root, leaf). Unidirectionally cloned. New library number
            given to library 707 for additional sequencing."

```

```

BASE COUNT      97 a  156 c  153 g  90 t
ORIGIN

Query Match      34.8%; Score 200; DB 10; Length 496;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 249 CACCACCACTATGTCACACGCGCGCGCGGCTTCGTCGCGAGACGAGCTTCCTCGCCACG 308
      |||||||
DB 183 CACCACCACTATGTCACACGCGCGCGGCTTCGTCGCGAGACGAGCTTCCTCGCCACG 242
      |||||||
QY 309 GAGGACATCAACACCTGCACCGCGCGGCTTCACGAGCGCAGGAGAGCTTCCTCGCCACG 368
      |||||||
DB 243 GAGGACATCAACACCTGCACCGCGCGGCTTCACGAGCGCAGGAGAGCTTCCTCGCCACG 302
      |||||||
QY 369 GCTAACTGAGCGCGCGCGCGCGGCTTCACGCGCGGCTTCGTCGCTGCGTCGCGCTTCA 428
      |||||||
DB 303 GCTAACTGAGCGCGCGCGCGCGGCTTCACGCGCGGCTTCGTCGCTGCGTCGCGCTTCA 362
      |||||||
QY 429 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
      |||||||
DB 363 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 422
      |||||||
QY 489 CACTCAGCTCCTGTACGAAATACGAAATACGAAATACGAAATACGAAATACGAAATACGAA 548
      |||||||
DB 423 CACTCAGCTCCTGTACGAAATACGAAATACGAAATACGAAATACGAAATACGAAATACGAA 482
      |||||||
QY 549 AT 550
      ||
DB 483 AT 484
      ||

```

```

RESULT 14
BQ619167.
LOCUS      648 bp mRNA linear EST 27-JUN-2002
DEFINITION RNOSEQ4E05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
mays cDNA clone RNOSEQ4E05_SK.ab1 similar to No homology, mRNA
sequence.
ACCESSION BQ619167
VERSION    1
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
REFERENCE  1 (bases 1 to 648)
AUTHORS    Wang,H. and Bohnert,H.J.
TITLE      Genomics of plant stress tolerance
JOURNAL    Unpublished (2002)
COMMENT    Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES   source
            1..648
            Location/Qualifiers
            1..648
            /organism="Zea mays"
            /db_xref="taxon:4577"
            /clone_lib="RNOSEQ4E05_SK.ab1"
            /clone_lib="Salt stressed Zea mays roots cDNA library"
            /tissue_type="Roots"
            /dev_stage="2 weeks old"
            /note="Vector: pBluescript SK+; Stressed 24 hours at 150
            mM NaCl"
BASE COUNT      172 a  193 c  183 g  100 t
ORIGIN

Query Match      31.0%; Score 178; DB 14; Length 648;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 328; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 244 ACCACCACACACCATGTCACACGCGCGCGCTTCGTGTCGCGAGACACGGG 303
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 233 ACCACCACACACCATGTCACACGCGCGCGCTTCGTGTCGCGAGACACGGG 292
 QY 304 TCGAAGAGGACATCAACACCTGCACGCGGAGGTCCACGAGCGGAGAGCTTCCTCG 363
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 293 TCGAAGAGGACATCAACACCTGCACGCGGAGGTCCACGAGCGGAGAGCTTCCTCG 352
 QY 364 CCAGGCTTAACCTGAGCCCGCGCGCGCGCATCCACGCCGCTTCGCTGCGGTG 423
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 353 CCAGGCTTAACCTGAGCCCGCGCGCGCGCATCCACGCCGCTTCGCTGCGGTG 412
 QY 424 CCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 413 CCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
 QY 484 GCAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 473 GCAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
 QY 544 TCGTAATACTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 574
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 533 TCGTAATACTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 563

RESULT 15

B0619315
 LOCUS
 DEFINITION RNOSE06C06_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays CDNA clone RNOSE06C06_SK.ab1 similar to No homology, mRNA sequence.
 ACCESSION B0619315
 VERSION B0619315.1 GI:21621309
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Wang, H. and Bohnert, H. J.
 TITLE Genomics of plant stress tolerance
 JOURNAL Unpublished (2002)
 COMMENT Contact: Mark Fredricksen
 Department of Plant Biology
 University of Illinois
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 2172655473
 Email: bohnertlab@life.uiuc.edu.

FEATURES

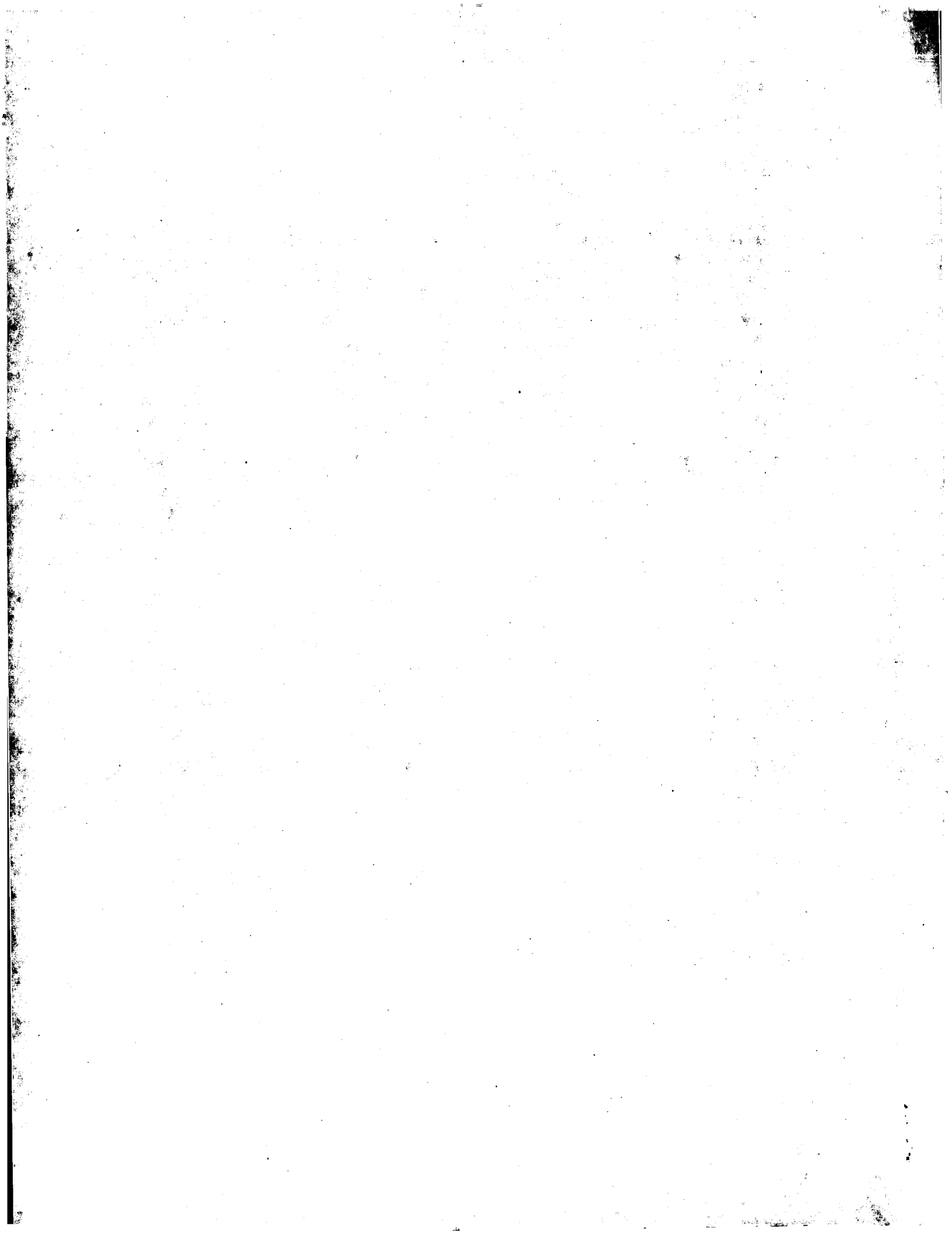
source
 1. 648
 Location/Qualifiers
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /clone="RNOSE06C06_SK.ab1"
 /clone_lib="Salt stressed Zea mays roots cDNA library"
 /tissue_type="Roots"
 /dev_stage="2 weeks old"
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
 BASE COUNT 172 a 193 c 183 g 100 t
 ORIGIN

Query Match 31.0%; Score 178; DB 14; Length 648;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 328; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 244 ACCACCACACACCATGTCACACGCGCGCGCTTCGTGTCGCGAGACACGGG 303
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 233 ACCACCACACACCATGTCACACGCGCGCGCTTCGTGTCGCGAGACACGGG 292
 QY 304 TCGAAGAGGACATCAACACCTGCACGCGGAGGTCCACGAGCGGAGAGCTTCCTCG 363
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 293 TCGAAGAGGACATCAACACCTGCACGCGGAGGTCCACGAGCGGAGAGCTTCCTCG 352

QY 364 CCAGGCTTAACCTGAGCCCGCGCGCGCTTCGTGTCGCGAGACACGGG 423
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 353 CCAGGCTTAACCTGAGCCCGCGCGCGCTTCGTGTCGCGAGACACGGG 412
 QY 424 CCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 413 CCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
 QY 484 GCAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 473 GCAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
 QY 544 TCGTAATACTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 574
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 533 TCGTAATACTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 563

Search completed: June 16, 2003, 08:38:46
 Job time : 1170.61 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 04:03:20 ; Search time 1984.2 Seconds
(without alignments)
9915.101 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 accacgcgtcgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_be.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	122	18.0	3454	9 HSM802993	AL713659 Homo sapi
2	121	17.9	596	9 BC008387	BC008387 Homo sapi
3	121	17.9	2776	10 BC034120	BC034120 Mus muscu
4	121	17.9	2830	9 HSM803477	AL832170 Homo sapi
5	121	17.9	3383	9 HSM803701	AL832393 Homo sapi
6	121	17.9	6486	6 AX458534	AX458534 Sequence
7	120	17.8	800	10 BC034163	BC034163 Mus muscu
8	120	17.8	1034	9 AK026865	AK026865 Homo sapi
9	120	17.8	1191	9 BC014547	BC014547 Homo sapi
10	120	17.8	1298	17 AF113689	AF113689 Homo sapi
11	120	17.8	1319	9 BC034528	BC034528 Homo sapi
12	120	17.8	1388	9 BC016332	BC016332 Homo sapi
13	120	17.8	1462	10 BC006018	BC006018 Mus muscu
14	120	17.8	1493	9 BC032462	BC032462 Homo sapi
15	120	17.8	1500	5 XLNAKAP	YL1387 X.laevis mr
16	120	17.8	1517	9 BC016152	BC016152 Homo sapi
17	120	17.8	1652	9 BC028131	BC028131 Homo sapi
18	120	17.8	1699	9 HSM801509	AL133640 Homo sapi
19	120	17.8	1728	9 AB070110	AB070110 Macaca fa
20	120	17.8	1802	9 BC025776	BC025776 Homo sapi
21	120	17.8	1953	9 AB056420	AB056420 Macaca fa
22	120	17.8	2003	17 AF130092	AF130092 Homo sapi
23	120	17.8	2074	9 BC014433	BC014433 Homo sapi
24	120	17.8	2082	6 AR059958	AR059958 Sequence
25	120	17.8	2105	9 HSM803705	AL832397 Homo sapi
26	120	17.8	2178	6 AX099370	AX099370 Sequence
27	120	17.8	2194	10 BC031202	BC031202 Mus muscu
28	120	17.8	2447	9 BC027927	BC027927 Homo sapi
29	120	17.8	2499	9 HSM801350	AL133016 Homo sapi
30	120	17.8	2833	9 HSM804701	AL833388 Homo sapi
31	120	17.8	2991	9 BC030556	BC030556 Homo sapi
32	120	17.8	3005	10 BC032271	BC032271 Mus muscu
33	120	17.8	4518	9 HSM802971	AL713745 Homo sapi
34	120	17.8	4930	9 HSM803373	AL832066 Homo sapi
35	120	17.8	5152	9 HSM803128	AL831816 Homo sapi
36	120	17.8	6171	6 AX345890	AX345890 Sequence
37	120	17.8	6577	6 AX251920	AX251920 Sequence
38	120	17.8	6577	6 AX346258	AX346258 Sequence
39	120	17.8	6577	6 AX348993	AX348993 Sequence
40	120	17.8	6794	6 AX251872	AX251872 Sequence
41	120	17.8	6794	6 AX344260	AX344260 Sequence
42	120	17.8	6794	6 AX348651	AX348651 Sequence
43	120	17.8	8899	6 AX345739	AX345739 Sequence
44	120	17.8	8899	6 AX348461	AX348461 Sequence
45	120	17.8	8946	6 AX345813	AX345813 Sequence

ALIGNMENTS

RESULT 1
HSM802993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

HSM802993
Homo sapiens mRNA; cDNA DKFZp667D1012 (from clone DKFZp667D1012).
AL713659
AL713659.1 GI:19584343
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 3454)
Koehler, K., Beyer, A., Mewes, H.W., Weill, B. and Wiemann, S.
Direct Submission
Submitted (12-MAR-2002) MIPS, Am Klopferspitze 18a, D-82152

COMMENT
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp667D1012) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
source
1. .3454
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="295.8 cR from top of Chr12 linkage group"
/clone="DKFZp667D1012"
/tissue_type="lymph node"
/clone_lib="667 (synonym: hlno2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
1. .3454
/gene="DKFZp667D1012"
1. .464
/gene="DKFZp667D1012"
/note="unknown zinc finger protein"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD28468.1"
/db_xref="GI:19584344"
/translation="HTGERPYSCLKCEKTFGRHHLIRHQTHLHDKTSRCSECGKNF RCNSHLASHORVHAEGKSCKGQGVESGTRKPRAPPVPVKCHVTECKSFGRHHL VRWLHTGKEKFPQCPRCSEKSGRKKHLDRHLLTHOGQSPRNSWDRTGSVF"
polyA_signal
3288. .3293
/gene="DKFZp667D1012"
3306
polyA_site
1058 a 711 c 756 g 929 t
BASE COUNT
ORIGIN
Query Match 18.0%; Score 122; DB 9; Length 3454;
Best Local Similarity 100.0%; Pred. No. 2.2e-55;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 555 CATCAA 614
DB 3306 CATCAA 3365
QY 615 AA 674
DB 3366 AA 3425
QY 675 AA 676
DB 3426 AA 3427
RESULT 2
LOCUS BC008387
DEFINITION Homo sapiens, clone IMAGE:4052341, mRNA.
ACCESSION BC008387
VERSION BC008387.1 GI:14249997
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 596)
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IMAGE Plate: 16 Row: P Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
FEATURES
Location/Qualifiers
1. 596
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4052341"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NIH_MGC_58"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
BASE COUNT 286 a 82 c 107 g 121 t
ORIGIN
Query Match 17.9%; Score 121; DB 9; Length 596;
Best Local Similarity 100.0%; Pred. No. 7.3e-55;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 556 ATCAA 615
DB 450 ATCAA 509
QY 616 AA 675
DB 510 AA 569
QY 676 A 676
DB 570 A 570
RESULT 3
LOCUS BC034120
DEFINITION Mus musculus, RIKEN cDNA 2810429K11 gene, clone MGC:31645
IMAGE:4527027, mRNA, complete cds.
ACCESSION BC034120
VERSION BC034120.1 GI:21707638
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2776)
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 44 Row: j Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

Location/Qualifiers

1..2776
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:31645 IMAGE:4527027"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 205..2157

CDS

/codon_start=1
 /product="RIKEN CDNA 2810429K11 gene"
 /protein_id="AAH34120.1"
 /db_xref="gi:21707639"
 /db_xref="locusID:74199"
 /translation="MGIVVPTMKASVIEVLVLVLTGIHNSKTPKTKRPKLTVPQI
 NCDVAKGIINPEFMVKPCAGQDPKYHVTGTVYASVSVCGAAIHSGVLDNSGGKI
 LVRKAGGSGYKSGVQSLPRWRESFVAESPKQGVAVPSLTLYSSSKTAA
 KAGETTKAYEKPISPTTIOPTLTQAQATPAEVRHSTSKPEASVTHSPRPQVG
 HRSMEEVDGKPGVLLDGSFVPEELSTQSEPPVQDPPCKDLDFLDGSTI
 GKRRFQKQFLADVQALDGPAGPLVGVVQDNPATQFNLTNHSODLKTAEK
 ITORGGLSNVGRALSFVITKFFSKANGRGAPNAVVMVDGPTDKVEVSRAVES
 GINVFPIVEGAARDIOHVVEPGFASKVCRTNGFYFNQVSLSHKTVQPLVRRV
 CDTDLACSTCLNSLDIGFVDSGMSGSTNFRTVQFVANSKEPEISDTOTRGA
 VOYTYEQLRSEFGDYNSKADILSAIRRVGWSGTTGAAIQVALBOLFKSKPNKR
 KVMIIITDGRSYDDVRIPAMAAYQKGVITVAIGIAWAQAQDELEVMTHPAKDISFFVD
 DFDNLKIAPRIQNICTEFNSQPN"
 BASE COUNT 831 a 627 c 683 g 635 t
 ORIGIN

Query Match 17.9%; Score 121; DB 10; Length 2776;
 Best Local Similarity 100.0%; Pred. No. 7.8e-55;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 556 ATCAAA 615
 DB 2629 ATCAAA 2688
 QY 616 AA 675
 DB 2689 AA 2748
 QY 676 A 676
 DB 2749 A 2749

RESULT 4

LOCUS HSM803477 HSM803477 2830 bp mRNA linear PRI 12-JUL-2002
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686E0316 (from clone DKFZp686E0316).
 ACCESSION AL832170
 VERSION AL832170.1 GI:21732714
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2830)
 Wambutt, R., Heubner, D., Mewes, H.W., Weill, B. and Wiemann, S.
 Direct Submission
 JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 consortium of the German Genome Project.
 This clone (DKFZp686E0316) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

Location/Qualifiers
 1..2830
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp686E0316"
 /tissue_type="cDNA-collection"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
 DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 polyA_signal 2686..2691
 polyA_site 2702
 BASE COUNT 982 a 438 c 583 g 827 t
 ORIGIN

Query Match 17.9%; Score 121; DB 9; Length 2830;
 Best Local Similarity 100.0%; Pred. No. 7.8e-55;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 556 ATCAAA 615
 DB 2706 ATCAAA 2765
 QY 616 AA 675
 DB 2766 AA 2825
 QY 676 A 676
 DB 2826 A 2826

RESULT 5

LOCUS HSM803701 HSM803701 3383 bp mRNA linear PRI 10-JUL-2002
 DEFINITION Homo sapiens mRNA; cDNA DKFZp667K1712 (from clone DKFZp667K1712).
 ACCESSION AL832393
 VERSION AL832393.1 GI:21732956
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3383)
 Koehrer, K., Beyer, A., Mewes, H.W., Weill, B. and Wiemann, S.
 Direct Submission
 JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the CDNA sequencing consortium of the German
 Genome Project.
 This clone (DKFZp667K1712) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

Location/Qualifiers
 1..3383
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone="DKFZp667K1712"		BC034163.1		GI:21706753	
/tissue_type="lymph node"		MGC.		house mouse.	
DH10B: sites NotI + SalI"		Mus musculus		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
/dev_stage="adult"		1 (bases 1 to 800)		Strausberg, R.	
3231..3236		Direct Submission		Submitted (02-JUL-2002)	
3250		Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		NIH-MGC Project URL: http://mgc.nci.nih.gov	
polyA_signal		1210 a		Contact: MGC help desk	
polyA_site		552 c		Email: cgapbs-remail.nih.gov	
BASE COUNT		552 c		CDNA Library Preparation: Life Technologies, Inc.	
ORIGIN		17.9%; Score 121; DB 9; Length 3383;		CDNA Library Arranged by: The I.M.A.G.E. Consortium (LLNL)	
Query Match		17.9%; Score 121; DB 9; Length 3383;		DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	
Best Local Similarity		100.0%; Pred. No. 7.9e-55;		Web site: http://www-shgc.stanford.edu	
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		556 ATCAAA		Contact: (Dickson, Mark) mcd@paxil.stanford.edu	
QY		556 ATCAAA		R. M.	
DB		3251 ATCAAA		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
QY		616 AA		Series: IRAK Plate: 58 Row: 0 Column: 24	
DB		3311 AA		This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753829.	
QY		676 A 676		Location/Qualifiers	
DB		3371 A 3371		1..800	
RESULT 6		AX458534/c		/organism="Mus musculus"	
LOCUS		AX458534		/db_xref="taxon:10090"	
DEFINITION		Sequence 80 from Patent WO0246454.		/map="CZCH II"	
ACCESSION		AX458534		/clone="MGC:36077 IMAGE:5065647"	
VERSION		AX458534.1		/tissue_type="Mammary tumor metastasized to lung, Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMIV."	
KEYWORDS		synthetic construct.		/clone_lib="NCI_CGAP_Lu29"	
SOURCE		synthetic construct		/lab_host="DH10B"	
ORGANISM		artificial sequences.		/note="Vector: pCMV-SPORT6"	
REFERENCE		1		45..305	
AUTHORS		Schacht, O.		/codon_start=1	
TITLE		Diagnosis of diseases associated with angiogenesis		/product="Fc receptor, IgE, high affinity I, gamma polypeptide"	
JOURNAL		Patent: WO 0246454-A 80 13-JUN-2002;		/protein_id="AAH34163.1"	
FEATURES		Epigenomics AG (DE)		/db_xref="GI:21706754"	
source		Location/Qualifiers		/db_xref="LocusID:14127"	
1..6486		/organism="synthetic construct"		/translation="MISAVILFLLLVFQAAALGEPOLCYILDVFLYGVITLLYC	
/db_xref="taxon:32630"		/note="chemically treated genomic DNA (Homo sapiens)"		RLKIQVRKAAIASREKADAVYTGILNRSQETETLKHEKPPQ"	
BASE COUNT		1713 a		BASE COUNT	
ORIGIN		176 c		ORIGIN	
Query Match		17.9%; Score 121; DB 6; Length 6486;		Query Match	
Best Local Similarity		100.0%; Pred. No. 8.1e-55;		Best Local Similarity	
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		556 ATCAAA		Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY		556 ATCAAA		QY	
DB		3483 ATCAAA		DB	
QY		616 AA		QY	
DB		3423 AA		DB	
QY		676 A 676		QY	
DB		3363 A 3363		DB	
RESULT 7		BC034163		RESULT 8	
LOCUS		BC034163		AK026865	
DEFINITION		Mus musculus, Fc receptor, IgE, high affinity I, gamma polypeptide, clone MGC:36077 IMAGE:5065647, mRNA, complete cds.		LOCUS	
ACCESSION		BC034163		DEFINITION	
				AF182844 Homo sapiens VPS28 protein mRNA.	

ACCESSION AK026865
 VERSION AK026865.1 GI:10439824
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens adipose tissue cDNA to mRNA, clone_lib:ad
 clone:ADS001518.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
 Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
 Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
 Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
 Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1034)
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing; Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source Location/Qualifiers
 1..1034
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADS001518"
 /tissue_type="adipose tissue"
 /clone_lib="ad"
 /note="cloning vector pME18SFL3"
 misc_feature 1..1034
 /note="highly similar to AF182844 Homo sapiens VPS28
 protein mRNA"
 BASE COUNT 315 a 289 c 264 g 166 t
 ORIGIN
 Query Match 17.8% Score 120; DB 9; Length 1034;
 Best Local Similarity 100.0%; Pred. No. 2.6e-54;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 557 TCAA 616
 Db 912 TCAA 971
 Qy 617 AA 676
 Db 972 AA 1031
 RESULT 9
 BC014547
 LOCUS BC014547 1191 bp mRNA linear PRI 26-SEP-2001
 DEFINITION Homo sapiens, similar to myoglobin, clone MGC:13548 IMAGE:4244838,
 mRNA, complete cds.
 ACCESSION BC014547
 VERSION BC014547.1 GI:15778932
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1191)
 AUTHORS Strausberg,R.

TITLE JOURNAL

REMARK COMMENT

Direct Submission
 Submitted (24-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amand@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 19 Row: 9 Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4885476.
 FEATURES
 source Location/Qualifiers
 1..1191
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:13548 IMAGE:4244838"
 /tissue_type="Skeletal Muscle"
 /clone_lib="NIH_MGC_81"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 77..541
 /codon_start=1
 /product="Similar to myoglobin"
 /protein_id="AAH14547.1"
 /db_xref="GI:15778933"
 /translation="MGLSDGKQLVNVKGVKVEADIPGHQVEVLRPKGHPETLEKFK
 DKFKHLKSEDEKAKSADLKKHGTATVLTGLKKKGHHEAEIKPLAQSHATKHKIPV
 KYLEFISECIQVLSKHPGDFGADGAMNKALELFKRKDMASNYKELGFGQ"
 BASE COUNT 357 a 278 c 314 g 242 t
 ORIGIN

Query Match 17.8% Score 120; DB 9; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 2.6e-54;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 557 TCAA 616
 Db 1072 TCAA 1131
 Qy 617 AA 676
 Db 1132 AA 1191
 RESULT 10
 AF113689
 ID AF113689 standard; RNA; HUM; 1298 BP.
 AC AF113689;
 XX AF113689;
 SV AF113689.1
 XX
 DT 03-FEB-2000 (Rel. 62, Created)
 DT 09-MAY-2001 (Rel. 67, Last updated, Version 2)
 XX
 DE Homo sapiens clone FLB4701.
 XX
 KW
 XX
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.

1

DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadansystemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 32 Row: D Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1. .1388
 /organism="Homo sapiens"
 /db_xref="LocusID:7336"
 /db_xref="taxon:9606"
 /clone="MGC:23763 IMAGE:4109228"
 /tissue_type="Bone marrow, acute myelogenous leukemia"
 /clone_lib="NIH_MGC_55"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

CDS

47. .484
 /codon_start=1
 /product="ubiquitin-conjugating enzyme E2 variant 2"
 /protein_id="AAH16332.1"
 /db_xref="GI:16740954"
 /translation="MAVSTGVKVPNFRLLLEELLEGKGVGDGTVSGLEDDEDMTIT
 RWCMILGPRTYENRIYLSKVECKYPEAPPSVRFVTKINNINGINSSGMVDARS
 IPVLAQWNSYSTRVQLQELRLMMSKRNKLPQPEGGQYNN"

BASE COUNT 592 a 179 c 271 g 346 t

Query Match 17.8% Score 120; DB 9; Length 1388;
 Best Local Similarity 100.0%; Pred. No. 2.7e-54;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAA 616

Db 1223 TCAA 1282

Qy 617 AA 676

Db 1283 AA 1342

RESULT 13
 BC006018

LOCUS Mus musculus, MyoD family inhibitor, clone MGC:5939 IMAGE:3488088,
 DEFINITION mRNA, complete cds.

ACCESSION BC006018.1 GI:13543726

VERSION MGC.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1462)
 Direct Submission
 Submitted (02-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk

COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 7 Row: O Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754665.

FEATURES

Location/Qualifiers

source

1. .1462

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="C57BL/6J"

/clone="MGC:5939 IMAGE:3488088"

/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months

old, gross tissue."

/clone_lib="NCI-CGAP_Mam5"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

26. .766

/codon_start=1

/product="MyoD family inhibitor"

/protein_id="AAH06018.1"

/db_xref="GI:13543727"

/translation="MSQVSGQCPSCDAPHGVPMSALDPAQTMSLLPGLGVARSTHPV

ESSESGFPEEAAPSPHDSGLRAQALNSIDLDPTEAVTCPOGPOGCTPLLPNG

SSHDHLSFEGSAGHACGALGGSKAHRKLTQHPSLGSOAGRSKRSASQVPLQAG

EDCCVHCILSCLECFELTLCNILLDCATCGSSSESDCLCCCGSGECADCDLPCLD

DGIVDACCESADCLCIECCGCGICFSS"

BASE COUNT 383 a 423 c 407 g 249 t

ORIGIN

Query Match 17.8% Score 120; DB 10; Length 1462;

Best Local Similarity 100.0%; Pred. No. 2.7e-54;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 557 TCAA 616

Db 1335 TCAA 1394

Qy 617 AA 676

Db 1395 AA 1454

RESULT 14
 BC032462

LOCUS Homo sapiens, hypothetical gene supported by BC017964, clone

DEFINITION MGC:40428 IMAGE:5197243, mRNA, complete cds.

ACCESSION BC032462

VERSION BC032462.1 GI:21595557

KEYWORDS MGC.

SOURCE human.

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 1493)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk

COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 03:58:29 ; Search time 222.269 seconds
(without alignments)
5849.143 Million cell updates/sec

Title: us-10-090-035-1
Perfect score: 676
Sequence: 1 accacgcgtccgcacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Genesec101002.*
1: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	121	17.9	381	23	ABV58662 Human prostate exp
2	121	17.9	1119	21	AAC60033 Human secreted pro
3	121	17.9	2645	22	AAS31262 Human cDNA encodin
4	121	17.9	2645	24	AB066586 Human polynucleoti
5	121	17.9	6486	24	AB067050 Human anglogenesis
6	120	17.8	439	23	ABV50206 Human prostate exp
7	120	17.8	452	23	ABV59282 Human prostate exp
8	120	17.8	484	23	ABV58008 Human prostate exp
9	120	17.8	516	23	ABV55889 Human prostate exp

10	120	17.8	544	23	ABV56624 Human prostate exp
11	120	17.8	579	23	ABV58690 DNA encoding human
12	120	17.8	804	22	AAS33148 Human secreted pro
13	120	17.8	1097	22	AAD08201 Human secreted pro
14	120	17.8	1097	22	AAD05430 DNA encoding novel
15	120	17.8	1194	23	ABK43496 Human polynucleoti
16	120	17.8	1204	22	AAI63823 Human secreted pro
17	120	17.8	1457	21	AAAG1287 Human secreted pro
18	120	17.8	1985	21	AAC59406 Mouse neuronal PAS
19	120	17.8	2082	19	AAV41257 Human cDNA clone B
20	120	17.8	2178	22	AAF98380 Human immune syste
21	120	17.8	6171	24	ABL32788 Chemically treated
22	120	17.8	6577	24	ABL70561 Human immune syste
23	120	17.8	6577	24	ABL33356 Human immune syste
24	120	17.8	6577	24	AAAG1221 Human gene regulat
25	120	17.8	6794	24	ABL70219 Chemically treated
26	120	17.8	6794	24	AAAG1174 Human gene regulat
27	120	17.8	6794	24	ABK31264 Signal transductio
28	120	17.8	8899	24	ABK40074 Human chemically p
29	120	17.8	8899	24	ABL32837 Human immune syste
30	120	17.8	8946	24	ABL32911 Chemically pretrea
31	120	17.8	15832	22	AAS45399 Human immune syste
32	120	17.8	15832	24	ABL33343 DNA transcription
33	120	17.8	15832	24	ABK28244 Human anglogenesis
34	120	17.8	40324	24	ABQ67149 Human nervous syst
35	119	17.6	165	22	ABA17044 Genomic sequence #
36	119	17.6	165	22	ABA08161 Human pancreatic f
37	119	17.6	165	22	AAS39812 Genomic sequence #
38	119	17.6	165	22	AAD16672 Human reproductive
39	119	17.6	165	22	AAS29235 Human reproductive
40	119	17.6	165	22	AAI04803 Human immune/haema
41	119	17.6	165	22	AAI07127 Human immune/haema
42	119	17.6	165	22	AAK73944 Human immune/haema
43	119	17.6	165	22	AAK78761 Human immune/haema
44	119	17.6	165	22	AAK79574 Human immune/haema
45	119	17.6	165	22	AAK81312 Human immune/haema

ALIGNMENTS

RESULT 1
ABV58662
ID ABV58662 standard; cDNA; 381 BP.
XX AC ABV58662;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 58653.
XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX

PN	WO200155368-A1.	
XX	02-AUG-2001.	
PD	17-JAN-2001; 2001WO-US01348.	
XX	31-JAN-2000; 2000US-0179065.	
PF	04-FEB-2000; 2000US-0180628.	
XX	24-FEB-2000; 2000US-0184664.	
	02-MAR-2000; 2000US-0186330.	
	16-MAR-2000; 2000US-0189874.	
	17-MAR-2000; 2000US-0190076.	
	18-APR-2000; 2000US-0198123.	
	19-MAY-2000; 2000US-0205515.	
	07-JUN-2000; 2000US-0209467.	
	28-JUN-2000; 2000US-0214886.	
	30-JUN-2000; 2000US-0215135.	
	07-JUL-2000; 2000US-0216647.	
	07-JUL-2000; 2000US-0216880.	
	11-JUL-2000; 2000US-0217487.	
	11-JUL-2000; 2000US-0217496.	
	14-JUL-2000; 2000US-0218290.	
	26-JUL-2000; 2000US-0220963.	
	26-JUL-2000; 2000US-0220964.	
	14-AUG-2000; 2000US-0224518.	
	14-AUG-2000; 2000US-0225213.	
	14-AUG-2000; 2000US-0225214.	
	14-AUG-2000; 2000US-0225266.	
	14-AUG-2000; 2000US-0225267.	
	14-AUG-2000; 2000US-0225268.	
	14-AUG-2000; 2000US-0225270.	
	14-AUG-2000; 2000US-0225447.	
	14-AUG-2000; 2000US-0225757.	
	14-AUG-2000; 2000US-0225758.	
	14-AUG-2000; 2000US-0225759.	
	18-AUG-2000; 2000US-0226279.	
	22-AUG-2000; 2000US-0226681.	
	22-AUG-2000; 2000US-0226868.	
	22-AUG-2000; 2000US-0227182.	
	23-AUG-2000; 2000US-0227009.	
	30-AUG-2000; 2000US-0228924.	
	01-SEP-2000; 2000US-0229247.	
	01-SEP-2000; 2000US-0229248.	
	01-SEP-2000; 2000US-0229343.	
	01-SEP-2000; 2000US-0229344.	
	01-SEP-2000; 2000US-0229345.	
	05-SEP-2000; 2000US-0229509.	
	05-SEP-2000; 2000US-0229513.	
	06-SEP-2000; 2000US-0230337.	
	06-SEP-2000; 2000US-0230438.	
	08-SEP-2000; 2000US-0231242.	
	08-SEP-2000; 2000US-0231243.	
	08-SEP-2000; 2000US-0231244.	
	08-SEP-2000; 2000US-0231413.	
	08-SEP-2000; 2000US-0231414.	
	08-SEP-2000; 2000US-0232080.	
	08-SEP-2000; 2000US-0232081.	
	12-SEP-2000; 2000US-0231968.	
	14-SEP-2000; 2000US-0232397.	
	14-SEP-2000; 2000US-0232398.	
	14-SEP-2000; 2000US-0232399.	
	14-SEP-2000; 2000US-0232400.	
	14-SEP-2000; 2000US-0232401.	
	14-SEP-2000; 2000US-0233063.	
	14-SEP-2000; 2000US-0233064.	
	21-SEP-2000; 2000US-0233065.	
	21-SEP-2000; 2000US-0234223.	
	21-SEP-2000; 2000US-0234274.	
	25-SEP-2000; 2000US-0234997.	
	25-SEP-2000; 2000US-0234998.	
	26-SEP-2000; 2000US-0235484.	
	27-SEP-2000; 2000US-0235834.	
	27-SEP-2000; 2000US-0235836.	
PR	29-SEP-2000; 2000US-0236327.	
PR	29-SEP-2000; 2000US-0236367.	
PR	29-SEP-2000; 2000US-0236368.	
PR	29-SEP-2000; 2000US-0236369.	
PR	29-SEP-2000; 2000US-0236370.	
PR	02-OCT-2000; 2000US-0236802.	
PR	02-OCT-2000; 2000US-0237037.	
PR	02-OCT-2000; 2000US-0237038.	
PR	02-OCT-2000; 2000US-0237039.	
PR	13-OCT-2000; 2000US-0237040.	
PR	13-OCT-2000; 2000US-0239335.	
PR	13-OCT-2000; 2000US-0239337.	
PR	20-OCT-2000; 2000US-0240960.	
PR	20-OCT-2000; 2000US-0241221.	
PR	20-OCT-2000; 2000US-0241785.	
PR	20-OCT-2000; 2000US-0241786.	
PR	20-OCT-2000; 2000US-0241787.	
PR	20-OCT-2000; 2000US-0241808.	
PR	20-OCT-2000; 2000US-0241809.	
PR	20-OCT-2000; 2000US-0241826.	
PR	01-NOV-2000; 2000US-0244617.	
PR	08-NOV-2000; 2000US-0246474.	
PR	08-NOV-2000; 2000US-0246475.	
PR	08-NOV-2000; 2000US-0246476.	
PR	08-NOV-2000; 2000US-0246477.	
PR	08-NOV-2000; 2000US-0246478.	
PR	08-NOV-2000; 2000US-0246523.	
PR	08-NOV-2000; 2000US-0246524.	
PR	08-NOV-2000; 2000US-0246525.	
PR	08-NOV-2000; 2000US-0246526.	
PR	08-NOV-2000; 2000US-0246527.	
PR	08-NOV-2000; 2000US-0246528.	
PR	08-NOV-2000; 2000US-0246532.	
PR	08-NOV-2000; 2000US-0246610.	
PR	08-NOV-2000; 2000US-0246611.	
PR	08-NOV-2000; 2000US-0246613.	
PR	17-NOV-2000; 2000US-0249207.	
PR	17-NOV-2000; 2000US-0249208.	
PR	17-NOV-2000; 2000US-0249209.	
PR	17-NOV-2000; 2000US-0249210.	
PR	17-NOV-2000; 2000US-0249211.	
PR	17-NOV-2000; 2000US-0249212.	
PR	17-NOV-2000; 2000US-0249213.	
PR	17-NOV-2000; 2000US-0249214.	
PR	17-NOV-2000; 2000US-0249215.	
PR	17-NOV-2000; 2000US-0249216.	
PR	17-NOV-2000; 2000US-0249217.	
PR	17-NOV-2000; 2000US-0249218.	
PR	17-NOV-2000; 2000US-0249244.	
PR	17-NOV-2000; 2000US-0249245.	
PR	17-NOV-2000; 2000US-0249264.	
PR	17-NOV-2000; 2000US-0249265.	
PR	17-NOV-2000; 2000US-0249297.	
PR	17-NOV-2000; 2000US-0249299.	
PR	01-DEC-2000; 2000US-0250160.	
PR	01-DEC-2000; 2000US-0250391.	
PR	05-DEC-2000; 2000US-0251030.	
PR	05-DEC-2000; 2000US-0251988.	
PR	05-DEC-2000; 2000US-0256719.	
PR	06-DEC-2000; 2000US-0251479.	
PR	08-DEC-2000; 2000US-0251856.	
PR	08-DEC-2000; 2000US-0251868.	
PR	08-DEC-2000; 2000US-0251869.	
PR	08-DEC-2000; 2000US-0251989.	
PR	08-DEC-2000; 2000US-0251990.	
PR	11-DEC-2000; 2000US-0254097.	
PR	05-JAN-2001; 2001US-0259678.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
XX		
PI		

XX	PD	11-APR-2002.	
XX	PF	17-JAN-2001;	2001US-0764870.
XX	PF	31-JAN-2000;	2000US-1790658P.
XX	PR	04-FEB-2000;	2000US-1790628P.
XX	PR	28-JUN-2000;	2000US-2148667P.
XX	PR	07-JUL-2000;	2000US-216647P.
XX	PR	07-JUL-2000;	2000US-216880P.
XX	PR	11-JUL-2000;	2000US-217487P.
XX	PR	11-JUL-2000;	2000US-217496P.
XX	PR	14-JUL-2000;	2000US-218290P.
XX	PR	26-JUL-2000;	2000US-2205963P.
XX	PR	26-JUL-2000;	2000US-220964P.
XX	PR	14-AUG-2000;	2000US-225418P.
XX	PR	14-AUG-2000;	2000US-225419P.
XX	PR	14-AUG-2000;	2000US-225267P.
XX	PR	14-AUG-2000;	2000US-225268P.
XX	PR	14-AUG-2000;	2000US-225270P.
XX	PR	14-AUG-2000;	2000US-225447P.
XX	PR	14-AUG-2000;	2000US-225457P.
XX	PR	14-AUG-2000;	2000US-225758P.
XX	PR	22-AUG-2000;	2000US-226668P.
XX	PR	30-AUG-2000;	2000US-228924P.
XX	PR	01-SEP-2000;	2000US-229287P.
XX	PR	01-SEP-2000;	2000US-229343P.
XX	PR	01-SEP-2000;	2000US-229344P.
XX	PR	01-SEP-2000;	2000US-229345P.
XX	PR	05-SEP-2000;	2000US-229509P.
XX	PR	05-SEP-2000;	2000US-229513P.
XX	PR	08-SEP-2000;	2000US-231413P.
XX	PR	21-SEP-2000;	2000US-234223P.
XX	PR	21-SEP-2000;	2000US-234274P.
XX	PR	25-SEP-2000;	2000US-234997P.
XX	PR	27-SEP-2000;	2000US-235834P.
XX	PR	29-SEP-2000;	2000US-236327P.
XX	PR	29-SEP-2000;	2000US-236367P.
XX	PR	29-SEP-2000;	2000US-236368P.
XX	PR	29-SEP-2000;	2000US-236369P.
XX	PR	29-SEP-2000;	2000US-236370P.
XX	PR	02-OCT-2000;	2000US-236802P.
XX	PR	02-OCT-2000;	2000US-237037P.
XX	PR	02-OCT-2000;	2000US-237038P.
XX	PR	02-OCT-2000;	2000US-237039P.
XX	PR	02-OCT-2000;	2000US-237040P.
XX	PR	13-OCT-2000;	2000US-237935P.
XX	PR	20-OCT-2000;	2000US-240960P.
XX	PR	20-OCT-2000;	2000US-241785P.
XX	PR	01-NOV-2000;	2000US-241809P.
XX	PR	01-NOV-2000;	2000US-244617P.
XX	PR	17-NOV-2000;	2000US-249239P.
XX	PR	08-DEC-2000;	2000US-251836P.
XX	PR	08-DEC-2000;	2000US-251866P.
XX	PR	08-DEC-2000;	2000US-251869P.
XX	PA	(ROSE/) ROSEN C A.	
XX	PA	(RUBE/) RUBEN S M.	
XX	PA	(BARA/) BARASH S C.	
XX	PI	Rosen CA, Ruben SM, Barash	
XX	XX	WPI: 2002-470713/50.	
XX	DR	P-PSDB; ABP47911.	
XX	XX	New nucleic acid encoding hu	
XX	PT	treatment and prevention of	
XX	PT	polypeptides and antibodies	
XX	PS	Claim 1; SEQ ID NO 76; 235p	
XX	XX	The invention relates to nov	
CC	CC	(ABP47846-ABP48110) useful	

CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870.

XX SQ Sequence 2645 BP; 827 A; 506 C; 513 G; 799 T; 0 other;

Query Match 17.9%; Score 121; DB 24; Length 2645;
 Best Local Similarity 100.0%; Pred. No. 1.7e-28;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 556 ATCAAA 615
 DB 2522 ATCAAA 2581
 QY 616 AA 675
 DB 2582 AA 2641
 QY 676 A 676
 DB 2642 A 2642

RESULT 5
 ABQ67050/c
 ID ABQ67050 standard; DNA; 6486 BP.
 XX AC ABQ67050;
 XX DT 28-AUG-2002 (first entry)
 XX DE Human angiogenesis associated polynucleotide SEQ ID NO 80.
 XX KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
 KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
 KW macular degeneration; inflammatory bowel disease; Crohn's disease;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiarteriosclerotic; ds.
 XX OS Homo sapiens.
 XX PN WO200246454-A2.
 XX PD 13-JUN-2002.
 XX PF 06-DEC-2001; 2001WO-EP14320.
 XX PR 06-DEC-2000; 2000DE-1061338.
 XX PA (EPIG-) EPIGENOMICS AG.
 XX PI Schacht O;
 XX DR WPI; 2002-500450/53.
 XX PT New nucleic acid fragments from chemically treated
 PT angiogenesis-associated genes, useful for determining methylation
 XX status, e.g. in diagnosis or treatment of cancer
 XX Claim 1; SEQ ID NO 80; 41pp + Sequence Listing; German.

XX The invention relates to a nucleic acid (I) comprising a segment of 18
 CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
 CC related oligomers, are used to evaluate the methylation status and/or
 CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
 CC diagnosis and treatment of eye diseases, proliferative retinopathy,
 CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 CC diabetic retinopathy, macular degeneration caused by neovascularisation,
 CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 CC Crohn's disease.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 6486 BP; 1713 A; 176 C; 1443 G; 3154 T; 0 other;

Query Match 17.9%; Score 121; DB 24; Length 6486;
 Best Local Similarity 100.0%; Pred. No. 1.3e-28;
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 556 ATCAAA 615
 DB 3483 ATCAAA 3424
 QY 616 AA 675
 DB 3423 AA 3364
 QY 676 A 676
 DB 3363 A 3363

RESULT 6
 ABV50206
 ID ABV50206 standard; cDNA; 439 BP.
 XX AC ABV50206;
 XX DT 17-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 50197.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX Claim 1; Page 9776; 11750pp; English.

```
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 439 BP; 179 A; 117 C; 69 G; 69 T; 5 other;
Query Match 17.8%; Score 120; DB 23; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.1e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 616
DB 113 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 172
QY 617 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
DB 173 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 232
RESULT 7
ABV59282
ID ABV59282 standard; cDNA; 452 BP.
XX
AC ABV59282;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 59273.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 11362; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 452 BP; 199 A; 91 C; 72 G; 89 T; 1 other;
Query Match 17.8%; Score 120; DB 23; Length 452;
Best Local Similarity 100.0%; Pred. No. 6e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 616
DB 243 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 302
QY 617 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
DB 303 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 362
RESULT 8
ABV58008
ID ABV58008 standard; cDNA; 484 BP.
XX
AC ABV58008;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 57999.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 11145; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
```

CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 484 BP; 230 A; 97 C; 82 G; 74 T; 1 other;

Query Match 17.8%; Score 120; DB 23; Length 484;
 Best Local Similarity 100.0%; Pred. No. 5.9e-28;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 TCAA 616
 Db 172 TCAA 231
 QY 617 AA 676
 Db 232 AA 291

RESULT 9
 ABV55889
 ID ABV55889 standard; cDNA; 516 BP.
 XX
 AC ABV55889;
 XX
 DT 17-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 55889.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 10790; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 516 BP; 239 A; 93 C; 91 G; 93 T; 0 other;

Query Match 17.8%; Score 120; DB 23; Length 516;
 Best Local Similarity 100.0%; Pred. No. 5.8e-28;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 TCAA 616
 Db 203 TCAA 262
 QY 617 AA 676
 Db 263 AA 322

RESULT 10
 ABV56624
 ID ABV56624 standard; cDNA; 544 BP.
 XX
 AC ABV56624;
 XX
 DT 17-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 56615.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 10918; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 544 BP; 220 A; 141 C; 97 G; 84 T; 2 other;
Query Match 17.8%; Score 120; DB 23; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.7e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
Db 123 TCAA 182
QY 617 AA 676
Db 183 AA 242
RESULT 11
ID ABV58690 standard; cDNA; 579 BP.
XX AC ABV58690;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 58681.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer
XX PS Claim 1; Page 11260-11261; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 579 BP; 269 A; 106 C; 82 G; 121 T; 1 other;
Query Match 17.8%; Score 120; DB 23; Length 579;
Best Local Similarity 100.0%; Pred. No. 5.6e-28;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
Db 168 TCAA 227
QY 617 AA 676
Db 228 AA 287
RESULT 12
AAS33148
ID AAS33148 standard; cDNA; 804 BP.
XX AC AAS33148;
XX DT 04-DEC-2001 (first entry)
XX DE DNA encoding human secreted protein, Seq ID No 107.
XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
XX KW Gaucher's disease; neurological disease; cerebrovascular disorder;
XX KW thrombosis; wound healing; ss.
XX OS Homo sapiens.
XX PN WO200155326-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01347.
XX PR 31-JAN-2000; 2000US-0179065.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
XX P-PSDB; AAU20439.
XX PT New nucleic acids and polypeptides, useful for diagnosing, preventing
XX PT or treating medical conditions -
XX PS Claim 1; SEQ ID No 107; 753pp; English.
XX CC The invention relates to novel isolated nucleic acid molecules (I)
XX CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
XX CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits, in
XX CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used
XX CC the prevention, treatment and diagnosis of diseases associated with
XX CC inappropriate expression of secreted proteins. (I) and complementary
XX CC sequences may also be used as DNA probes in diagnostic assays (e.g.
XX CC polynucleotide chain reactions (PCR)) to detect and quantitate the presence
XX CC of similar nucleic acid sequences in samples, and so which patients may
XX CC be in need of restorative therapy. (II) may also be used as antigens in
XX CC the production of antibodies and in assays to identify modulators

(agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease and Charcot-Marie-Tooth disease), e.g. disease, cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and thrombosis), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). (I) and (II), agonists, antagonists and antibodies can also be used to promote wound healing, maintain organs before transplantation, and support cell culture of primary tissues. AHS33043-AHS33486 represent human secreted protein coding sequences, PCR primers, and related sequences of the invention. Note: the sequence data for this patent did not appear in the printed specification but was obtained in electronic format directly from WIPO at: [ftp.wipo.int/pub/published_pct_sequences](http://www.int/pub/published_pct_sequences).

Sequence 804 BP; 286 A; 167 C; 208 G; 143 T; 0 other;

Query Match	17.88;	Score 120;	DB 22;	Length 804;
Best Local Similarity	100.08;	Pred. No. 5e-28;		

Accession	Protein	Length	Score	E-value
557	TC	616	616	616
677	TC	736	736	736
617	QA	676	676	676
737	QA	796	796	796

RESULT 13

AAD08201
 ID AAD08201 standard; cDNA; 1097 BP.

AC AAD08201;

DT 07-AUG-2001 (first entry)

Human secreted protein-encoding gene 11 cDNA clone HCNDV12, SEQ ID NO:21.

Human; secreted protein; proliferative disorder; cancer; tumour; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiotenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; endocrine disorder; infection; wound healing; vulnery; cell culture; chemotaxis; food additive; gene therapy; binding partner identification; ss.

OS Homo sapiens.

2A	Key	Location/Qualifiers
FH	CDS	193..378
FT		

```

Z1
FT
      /*tag= a
      /product= "Human secreted protein precursor"

```

ET	sig_peptide
	193..264

ET / *tag= b

ET	mat_peptide	265..375	/*tag= C
ET			

PN WO200132837-A1.

PD 10-MAY-2001.
XX
XX PF 17-OCT-2000; 2000WO-US38664.
XX
XX PR 02-NOV-1999; 99US-0163085.
PR 17-DEC-1999; 99US-0172411.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Fiscella M, Wei P, Lafleur DW, Olsen HS, Baker K, Ebner R;
PI Komatsoulis G, Rosen CA, Ruben SM, Duan RD, Young PE, Florence KA;
PI Moore PA, Birse CE, Ni J, Soppet DR, Shi Y;
XX
DR WPI: 2001-328782/34.
DR P-PSDB: AAE03774.
XX
XX Novel human secreted proteins and nucleic acids for diagnosing,
PT preventing and treating neurological, cardiovascular, infectious,
PT autoimmune, gastrointestinal, bone disorders, cancer, particularly
PT ovarian cancer -
YV

AA080191-AAD08213 represent cDNAs corresponding to 19 human secreted protein genes, and AA03764-AA03786 represent the proteins they encode. AA03787-AA03800 represent human secreted protein fragments or variants. The secreted proteins and their genes are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 19 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS; autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiotensin disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein-encoding cDNA of the invention.

SQ Sequence 1097 BP; 458 A; 182 C; 231 G; 226 T; 0 other;

Query Match	17.8%;	Score 120;	DB 22;	Length 1097;
Best Local Similarity	100.0%;	Pred. NO;	4.5e-28;	
Matches 120: Conservative	0;	Mismatches	0;	Indels

Qy	557	TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	616
Dd	977	TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1036
Qy	617	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	676
Dd	1037	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1096

RESULT 14

AAD05430
 ID AAD05430 standard; cDNA; 1097 bp

AA AAD05430;

AA AAD05430;

PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226888.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-581633/65.

P-PSDB; AAU87166.

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -

Claim 1; SEQ ID No 86; 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (II) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial

CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 17.8%; Score 120; DB 23; Length 1194;
 Best Local Similarity 100.0%; Pred. No. 4.4e-28;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 TCACAAAAA 616
 Db 1070 TCACAAAAA 1129
 QY 617 AAAAAAA 676
 Db 1130 AAAAAAA 1189

Search completed: June 16, 2003, 06:54:02
 Job time : 230.769 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 06:35:31 ; Search time 47.0496 Seconds
(without alignments)
4406.275 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 acccacggtccgccacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	17.8	2082	2	US-08-785-310A-2
2	119	17.6	3275	4	US-09-370-838-151
3	118	17.5	144	1	US-08-702-344-26
4	118	17.5	240	1	US-08-628-417-6
5	118	17.5	1051	4	US-09-245-041-10
6	118	17.5	1447	4	US-09-443-041A-27
7	118	17.5	1798	4	US-09-797-906-1
8	118	17.5	2447	2	US-09-014-969-14
9	118	17.5	2674	4	US-09-817-180-1
10	117	17.3	117	1	US-08-702-344-3
11	117	17.3	1733	3	US-09-073-569-1
12	116	17.2	1813	3	US-09-071-224-3
13	114	16.9	2269	4	US-09-394-645-1
14	114	16.9	2269	4	US-09-243-560B-1
15	113	16.7	2246	4	US-09-363-708-3
16	111	16.4	2634	4	US-09-463-238-3
17	109	16.1	111	4	US-09-297-535-23
18	109	16.1	121	4	US-09-297-535-20
19	109	16.1	2186	4	US-09-360-545-66
20	108	16.0	9589	1	US-07-925-695-1
21	108	16.0	9589	1	US-07-925-695-2
22	106	15.7	2184	4	US-08-955-918C-1
23	106	15.7	2184	4	US-08-697-766A-1
24	105	15.5	1872	4	US-09-801-052-1
25	104	15.2	1066	1	US-08-157-101A-4
26	103	15.2	1882	4	US-09-370-253-1
27	101	14.9	140	1	US-08-628-417-5

28 100 14.8 578 4 US-09-602-877A-95
29 99 14.6 1474 4 US-08-821-994-64
30 99 14.6 6671 1 US-08-280-443-1
31 99 14.6 6671 1 US-08-457-459-1
32 99 14.6 6671 1 US-08-535-678-1
33 99 14.6 6671 5 PCT-US95-02275-1
34 98 14.5 98 1 US-08-088-658-42
35 98 14.5 98 2 US-08-471-907A-42
36 98 14.5 1454 4 US-09-372-422A-19
37 96 14.2 1117 4 US-09-247-373B-33
38 95 14.1 2806 4 US-09-653-839-9
39 93 13.8 1098 3 US-09-248-335-35
40 92 13.6 1582 3 US-08-545-196B-10
41 92 13.6 1582 3 US-08-545-196B-12
42 91 13.5 2323 4 US-09-149-476-24
43 91 13.5 2671 6 5168051-9
44 90 13.3 2665 4 US-08-971-089-5
45 89 13.2 1193 4 US-09-372-422A-23

ALIGNMENTS

RESULT 1
US-08-785-310A-2
; Sequence 2, Application US/08785310A
; Patent No. 5840532
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; TITLE OF INVENTION: Neuronal PAS Domain Protein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,310A
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-785-310A-2

Query Match 17.8%; Score 120; DB 2; Length 2082;
Best Local Similarity 100.0%; Pred. No. 1.8e-35;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAA 616
DB 1958 TCAA 2017
QY 617 AA 676

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 559 AA 618
| | | | |
Db 40 AA 99
| | | | |
Qy 619 AA 676
| | | | |
Db 100 AA 157
| | | | |

RESULT 5

US-09-245-041-10
; Sequence 10, Application US/09245041

; Patent No. 6274339

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT

; FILE REFERENCE: 7853-136

; CURRENT APPLICATION NUMBER: US/09/245.041

; EARLIER FILING DATE: 1998-07-21

; EARLIER APPLICATION NUMBER: 60/093,630

; EARLIER FILING DATE: 1998-07-21

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 1051

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-245-041-10

Query Match 17.5%; Score 118; DB 4; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 559 AA 618
| | | | |
Db 10 AA 69
| | | | |
Qy 619 AA 676
| | | | |
Db 70 AA 127
| | | | |

RESULT 6

US-09-443-041A-27
; Sequence 27, Application US/09443041A

; Patent No. 6465717

; GENERAL INFORMATION:

; APPLICANT: Farnodu, Omolayo O.

; APPLICANT: Orozco, Buddy

; APPLICANT: Rafalski, Antoni

; APPLICANT: Shen, Jennie

; TITLE OF INVENTION: Sterol Metabolism Enzymes

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/09/443,041A

; PRIOR FILING DATE: 1999-11-18

; PRIOR APPLICATION NUMBER: 60/109,283

; PRIOR FILING DATE: 1998-11-20

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 27

; LENGTH: 1447

; TYPE: DNA

; ORGANISM: Glycine max

US-09-443-041A-27

Query Match 17.5%; Score 118; DB 4; Length 1447;
Best Local Similarity 100.0%; Pred. No. 1e-34;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 559 AA 618
| | | | |
Db 1318 AA 1377
| | | | |
Qy 619 AA 676
| | | | |
Db 1378 AA 1435
| | | | |

RESULT 7

US-09-797-906-1

; Sequence 1, Application US/09797906

; Patent No. 6329188

; GENERAL INFORMATION:

; APPLICANT: Zlanghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; FILE REFERENCE: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; CURRENT APPLICATION NUMBER: US/09/797,906

; CURRENT FILING DATE: 2001-03-05

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1798

; TYPE: DNA

; ORGANISM: Human

US-09-797-906-1

Query Match 17.5%; Score 118; DB 4; Length 1798;

Best Local Similarity 100.0%; Pred. No. 9.9e-35;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 559 AA 618
| | | | |
Db 1645 AA 1704
| | | | |
Qy 619 AA 676
| | | | |
Db 1705 AA 1762
| | | | |

RESULT 8

US-09-014-969-14

; Sequence 14, Application US/09014969

; Patent No. 5965397

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth

; APPLICANT: McCoy, John M.

; APPLICANT: Lavallee, Edward R.

; APPLICANT: Racie, Lisa A.

; APPLICANT: Merberg, David

; APPLICANT: Treacy, Maurice

; APPLICANT: Spaulding, Vikki

; APPLICANT: Agostino, Michael J.

; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESS: Genetics Institute, Inc.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: MA

; COUNTRY: U.S.A.

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-014-969-14

Query Match 17.5%; Score 118; DB 2; Length 2447;
Best Local Similarity 100.0%; Pred. No. 9.3e-35;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 2260 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2319

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
Db 2320 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2377

RESULT 9
US-09-817-180-1
; Sequence 1, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CL001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
; US-09-817-180-1

Query Match 17.5%; Score 118; DB 4; Length 2674;
Best Local Similarity 100.0%; Pred. No. 9.1e-35;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 2547 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2606

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 676
Db 2607 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2664

RESULT 10
US-08-702-344-3
; Sequence 3, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa

; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-702-344-3

Query Match 17.3%; Score 117; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.1e-34;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 618
Db 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

QY 619 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 675
Db 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 117

RESULT 11
US-09-073-569-1
; Sequence 1, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
```

;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sawislak, Deborah A
;; REGISTRATION NUMBER: 37,438
;; REFERENCE/DOCKET NUMBER: 97-14
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6672
;; TELEFAX: 206-442-6678
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1733 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: Coding Sequence
;; LOCATION: 34...1344
;; OTHER INFORMATION:
US-09-073-569-1

Query Match 17.3%; Score 117; DB 3; Length 1733;
Best Local Similarity 100.0%; Pred. No. 2.3e-34;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 CAAAAA...AA 617
DB 1607 CAAAAA...AA 1666

QY 618 AAAAAA...AA 674
DB 1667 AAAAAA...AA 1723

RESULT 12
US-09-071-224-3
; Sequence 3, Application US/09071224
; Patent No. 6271343
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Presnell, Scott R.
; APPLICANT: Jelmeberg, Anna C.
; APPLICANT: Gilbert, Teresa
; APPLICANT: Foster, Donald C.
; APPLICANT: Adams, Robyn L.
; APPLICANT: Lehner, Joyce M.
; TITLE OF INVENTION: MAMMALIAN ZCYTORS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics
; STREET: 1201 Eastlake Ave East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,224
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G

;; REGISTRATION NUMBER: 32,743
;; REFERENCE/DOCKET NUMBER: 96-22
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6627
;; TELEFAX: 206-442-6678
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1813 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: Coding Sequence
;; LOCATION: 88...1362
;; OTHER INFORMATION:
US-09-071-224-3

Query Match 17.2%; Score 116; DB 4; Length 1813;
Best Local Similarity 100.0%; Pred. No. 5.2e-34;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 CAAAAA...AA 617
DB 1687 CAAAAA...AA 1746

QY 618 AAAAAA...AA 673
DB 1747 AAAAAA...AA 1802

RESULT 13
US-09-394-645-1
; Sequence 1, Application US/09394645
; Patent No. 6380371
; GENERAL INFORMATION:
; APPLICANT: Sassetti, Christopher M.
; APPLICANT: Rosen, Steven R.
; TITLE OF INVENTION: Endoglycan: A NOVEL PROTEIN HAVING SELECTIN
; TITLE OF INVENTION: LIGAND AND CHEMOKINE PRESENTATION ACTIVITY
; FILE REFERENCE: 6510-122US1
; CURRENT APPLICATION NUMBER: US/09/394,645
; CURRENT FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: human
US-09-394-645-1

Query Match 16.9%; Score 114; DB 4; Length 2269;
Best Local Similarity 100.0%; Pred. No. 2.6e-33;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 559 AAAAAA...AA 618
DB 2156 AAAAAA...AA 2215

QY 619 AAAAAA...AA 672
DB 2216 AAAAAA...AA 2269

RESULT 14
US-09-243-560B-1
; Sequence 1, Application US/09243560B
; Patent No. 6395882
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sassetti, Christopher M.
; TITLE OF INVENTION: No. 6395882el Selectin Ligands

FILE REFERENCE: UCAL097US1
CURRENT APPLICATION NUMBER: US/09/243,560B
CURRENT FILING DATE: 1999-02-03
PRIOR APPLICATION NUMBER: US 60/074,389
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2269
TYPE: DNA
ORGANISM: Homo sapiens
US-09-243-560B-1

Query Match 16.7%; Score 114; DB 4; Length 2269;
Best Local Similarity 100.0%; Pred. No. 2.6e-33;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 559 AA 618
DB 2156 AA 2215
QY 619 AA 672
DB 2216 AA 2269

RESULT 15
US-09-363-708-3
Sequence 3, Application US/09363708
Patent No. 6399747
GENERAL INFORMATION:
APPLICANT: Schmandt, et al.
TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/363,708
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/34451
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2246 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
DESCRIPTION: /desc = "mouse PAL cdna"
US-09-363-708-3

Query Match 16.7%; Score 113; DB 4; Length 2246;
Best Local Similarity 100.0%; Pred. No. 6.1e-33;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 558 CAAA 617

Db 2134 CAAA 2193
QY 618 AA 670
DB 2194 AA 2246
Search completed: June 16, 2003, 08:47:01
Job time : 54.0496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2003, 07:57:31 ; Search time 149.802 Seconds
(without alignments)
6534.570 Million cell updates/sec

Title: US-10-090-035-1
Perfect score: 676
Sequence: 1 accacgcgctgcgcacgc.....aaaaaaaaaaaaaaaaaaaaa 676

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1029858 seqs, 724030393 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	676	100.0	676	9	US-10-090-035-1
2	313	46.3	574	9	US-10-090-035-3
3	313	46.3	577	9	US-10-090-035-5
4	261	38.6	580	9	US-10-090-035-7
5	188	27.8	529	9	US-10-090-035-9
6	169	25.0	524	9	US-10-090-035-17
7	121	17.9	2645	9	US-10-125-540-76
8	121	17.9	2645	10	US-09-764-870-76
9	120	17.8	1097	9	US-09-832-129-21
10	120	17.8	1457	9	US-09-948-820-38
c 11	120	17.8	15832	9	US-10-239-676-106
c 12	119	17.6	160	10	US-09-960-352-11058
13	119	17.6	165	9	US-09-764-872-709
c 14	119	17.6	165	9	US-10-091-483-344
15	119	17.6	165	9	US-09-764-891-7491
16	119	17.6	165	9	US-09-764-891-9815
c 17	119	17.6	165	10	US-09-764-846-344
c 18	119	17.6	181	10	US-09-960-352-1301
c 19	119	17.6	217	10	US-09-960-352-6539
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 9, Appli
					Sequence 17, Appli
					Sequence 76, Appli
					Sequence 21, Appli
					Sequence 38, Appli
					Sequence 106, Appli
					Sequence 11058, A
					Sequence 709, Appli
					Sequence 344, Appli
					Sequence 7491, Appli
					Sequence 9815, Appli
					Sequence 344, Appli
					Sequence 1301, Appli
					Sequence 6539, Appli

20	119	17.6	409	10	US-09-822-849A-37	Sequence 37, Appli
21	119	17.6	495	10	US-09-878-574-4765	Sequence 4765, Appli
22	119	17.6	569	9	US-10-012-542-44	Sequence 44, Appli
23	119	17.6	674	9	US-10-076-622-465	Sequence 465, Appli
24	119	17.6	674	9	US-09-551-621-465	Sequence 21, Appli
25	119	17.6	674	10	US-09-825-301-21	Sequence 465, Appli
26	119	17.6	674	10	US-09-604-287A-465	Sequence 465, Appli
27	119	17.6	674	12	US-10-007-805-465	Sequence 26, Appli
c 28	119	17.6	681	10	US-09-822-830A-26	Sequence 161, Appli
29	119	17.6	758	10	US-09-764-853-161	Sequence 67, Appli
30	119	17.6	831	10	US-09-729-674-67	Sequence 212, Appli
31	119	17.6	1985	9	US-09-905-291A-212	Sequence 212, Appli
32	119	17.6	1985	9	US-09-902-853-212	Sequence 212, Appli
33	119	17.6	1985	9	US-09-907-824-212	Sequence 212, Appli
34	119	17.6	1985	9	US-09-907-841-212	Sequence 212, Appli
35	119	17.6	1985	9	US-09-904-011-212	Sequence 212, Appli
36	119	17.6	1985	9	US-10-174-590-27	Sequence 27, Appli
37	119	17.6	1985	9	US-10-176-758-27	Sequence 27, Appli
38	119	17.6	1985	9	US-10-175-737-27	Sequence 27, Appli
39	119	17.6	1985	9	US-09-906-742-212	Sequence 212, Appli
40	119	17.6	1985	9	US-10-173-706-27	Sequence 27, Appli
41	119	17.6	1985	9	US-10-175-738-27	Sequence 27, Appli
42	119	17.6	1985	9	US-10-175-752-27	Sequence 27, Appli
43	119	17.6	1985	9	US-10-176-482-27	Sequence 27, Appli
44	119	17.6	1985	9	US-10-176-757-27	Sequence 27, Appli
45	119	17.6	1985	9	US-10-176-913-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1
US-10-090-035-1
; Sequence 1, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: SIMMONS, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(367)
US-10-090-035-1

Query Match	100.0%	Score 676;	DB 9;	Length 676;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 676;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACCCACGCGTCCGCCACCGGTC	CGGTC	CGGTC
DB	1	ACCCACGCGTCCGCCACCGGTC	CGGTC	CGGTC
QY	61	TCGGAGGACACACCAAGCGT	TCGCA	TCGCA
DB	61	TCGGAGGACACACCAAGCGT	TCGCA	TCGCA
QY	121	GGAGGAGGTGAGTTCGGTGG	CGCGGTC	CGCGGTC
DB	121	GGAGGAGGTGAGTTCGGTGG	CGCGGTC	CGCGGTC
QY	181	GCACGTCGTCAAGGAGAAAGT	TCGAG	TCGAG
DB	181	GCACGTCGTCAAGGAGAAAGT	TCGAG	TCGAG

Db	181	GCACGTCGTCAAGGAGAGTTCCAGAGAGTCCACACGGTATACACCGCGCGGCCCAACCA	240
QY	241	CCACCACCACTATGTCACACCGCGGCCACCGCTTCGTGGTCCCGAGACACAGGTTGCA	300
Db	241	CCACCACCACTATGTCACACCGCGGCCACCGCTTCGTGGTCCCGAGACACAGGTTGCA	300
QY	301	GGAGGACATCAACACCTGACACCGCGGAGGTCCACGAGCGCAGGAGGCTTCTCGCCAG	360
Db	301	GGAGGACATCAACACCTGACACCGCGGAGGTCCACGAGCGCAGGAGGCTTCTCGCCAG	360
QY	361	GGTAACTCAGCGCGCGCGCGCATCCACGCGCGCTTCGTTCGCTCGCGCTT	420
Db	361	GGTAACTCAGCGCGCGCGCGCATCCACGCGCGCTTCGTTCGCTCGCGCTT	420
QY	421	ATGATATGCTGTGGTGTACTGTTGTGAGGTCATCGTACTTGGCTATCGTACGTGCAC	480
Db	421	ATGATATGCTGTGGTGTACTGTTGTGAGGTCATCGTACTTGGCTATCGTACGTGCAC	480
QY	481	GCACTCAGTCCCTGTACGAATACGACAAATAGCTCGTGACCTGAATAAACTTCTTCGT	540
Db	481	GCACTCAGTCCCTGTACGAATACGACAAATAGCTCGTGACCTGAATAAACTTCTTCGT	540
QY	541	AATACATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	600
Db	541	AATACATACCTACATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	600
QY	601	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	660
Db	601	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	660
QY	661	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	720
Db	661	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	720
RESULT 2			
US-10-090-035-3			
; Sequence 3, Application US/10090035			
; Patent No. US20020170089A1			
; GENERAL INFORMATION:			
; APPLICANT: Simmons, Carl R.			
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible			
; FILE REFERENCE: 35718/242990			
; CURRENT APPLICATION NUMBER: US/10/090,035			
; PRIOR FILING DATE: 2002-02-28			
; PRIOR APPLICATION NUMBER: 60/272,227			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 574			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (96)...(374)			
US-10-090-035-3			
Query Match			
Best Local Similarity 46.3%; Score 313; DB 9; Length 574;			
Matches 463; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	83	GCACCAATGGCTTACTACACGAGGTGAGTCTGCTCGGAGAGGTGAGGTGCGTGGCC	142
Db	90	GCACCAATGGCTTACTACACGAGGTGAGTCTGCTCGGAGAGGTGAGGTGCGTGGCC	149
QY	143	CCGCGCGGCTTCGGCGCCACCGCGGCGTCCAGCAGACGTCTGTCAGGAGAGTTC	202
Db	150	CCGCGCGGCTTCGGCGCCACCGCGGCGTCCAGCAGACGTCTGTCAGGAGAGTTC	209
QY	203	GAGGAGGTCCACAGGTATACGCGCGCGCGCCCAACACCACTACCTGTCACAC	262
US-10-090-035-5			
Query Match			
Best Local Similarity 46.3%; Score 313; DB 9; Length 577;			
Matches 463; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	83	GCACCAATGGCTTACTACACGAGGTGAGTCTGCTCGGAGAGGTGAGGTGCGTGGCC	142
Db	93	GCACCAATGGCTTACTACACGAGGTGAGTCTGCTCGGAGAGGTGAGGTGCGTGGCC	152
QY	143	CCGCGCGGCTTCGGCGCCACCGCGGCGTCCAGCAGACGTCTGTCAGGAGAGTTC	202
Db	153	CCGCGCGGCTTCGGCGCCACCGCGGCGTCCAGCAGACGTCTGTCAGGAGAGTTC	212
QY	203	GAGGAGGTCCACAGGTATACGCGCGCGCGCCCAACACCACTACCTGTCACAC	262
Db	213	GAGGAGGTCCACAGGTATACGCGCGCGCGCCCAACACCACTACCTGTCACAC	272
QY	263	GGCGGCCACGGCTTCGTGGTGGCGGAGACGAGGTGCGGAGGAGATCAACACCTCAC	322
Db	273	GGCGGCCACGGCTTCGTGGTGGCGGAGACGAGGTGCGGAGGAGATCAACACCTCAC	332
QY	323	GGCGAGGTCCACAGGCGGAGGAGGCTTCTCGCCAGGCTTAATGAGCGCGCGCGGG	382
Db	333	GGCGAGGTCCACAGGCGGAGGAGGCTTCTCGCCAGGCTTAATGAGCGCGCGCGGG	392
QY	383	CCGCGCATCCACGCGCGGCTTCGTGGTGGCTGCTTATGTTATGTTGTTGACTGG	442

393	CGGCATCCAGCCCGGCTTCGCTGCTCGGTGCTTATGATATCTGTGTTGACTCGG	Db
443	TTGTGCAGGGTCATCGTACTTTGGCTATCGTACGTGCACGCACCTCAGCTCCTCTACGAATT	Qy
453	TTGTTCAGGGTCATCGTACTTTGGGTATCGTACGTGCACGCACCTCAGCTCCTCTACGAATT	Db
503	ACGACAATAAGCTCGTGACCTCGAATAAAACTTCTTCGTAATACTAA	Qy
513	ACGACAATAAGCTCGTGACCTCGAATAAAACTTCTTCGTAATACTAA	Db
502		
512		
548		
558		

```

RESULT 4
US-10-090-035-7
; Sequence 7, Application US/10090035
; Patent No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)...(380)
US-10-090-035-7

```

RESULT 5
US-10-090-035-9
: Sequence 9, Application US/10090035
: Patent No. US20020170089A1
: GENERAL INFORMATION:
: APPLICANT: Simmons, Carl R.
: TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
: TITLE OF INVENTION: Proteins and Uses Thereof
: FILE REFERENCE: 35718/242990

```

; CURRENT APPLICATION NUMBER: US/10/090_035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(331)
US-10-090-035-9

Query Match      27.8%; Score 188; DB 9; Length 529;
Best Local Similarity 99.3%; Pred. No. 1.7e-82;
Matches 288; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      83 GCACCAATGGCTTACTACAGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGTGGCC 142
Db      47 GCACCAATGGCTTACTACAGAGGAGGTGGACTACTGCTCGGAGGAGGTGAGGTGCGTGGCC 106
Qy      143 CCGGCCGGCTTCGGCCGCCACGGGGGGGGTCCAGCAGCAGCGTCGTCAAGGAGAAAGTTC 202
Db      107 CCGGCCGGCTTCGGCCGCCACGGGGGGGGTCCAGCAGCAGCGTCGTCAAGGAGAAAGTTC 166
Qy      203 GAGGAGGTGCACACGGTATCACGGCGCGGGGCCCAACCACCACCACCATGGTTCACCAC 262
Db      167 GAGGAGGTGCACACGGTTCGACACGGCGCGGGGCCCAACCACCACCACCATGGTTCACCAC 226
Qy      263 GGGCGGCCACCGCTTCGTGGTTCGGCGACACAGGGTCGAGGAGGACATCAACACCTGCACC 322
Db      227 GGGCGGCCACCGCTTCGTGGTTCGGCGACACAGGGTCGAGGAGGACATCAACACCTGCACC 286
Qy      323 GGGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCACGGGTAACTGAGC 372
Db      287 GGGGAGGTCCACGAGCGCAGGAGAGGTTCCTCGCCACGGGTAACTGAGC 336

```

```

RESULT 6
US-10-090-035-17
: Sequence 17, Application US/10090035
: Patent No. US20020170089A1
: GENERAL INFORMATION:
: APPLICANT: Simmons, Carl R.
: TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
: TITLE OF INVENTION: Proteins and Uses Thereof
: FILE REFERENCE: 35718/242990
: CURRENT APPLICATION NUMBER: US/10/090.035
: CURRENT FILING DATE: 2002-02-28
: PRIOR APPLICATION NUMBER: 60/272,227
: PRIOR FILING DATE: 02/28/2001
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 524
: TYPE: DNA
: ORGANISM: Triticum aestivum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (57)...(338)
: NAME/KEY: misc_feature
: LOCATION: (1)...(524)
: OTHER INFORMATION: n = A,T,C or G
US-10-090-035-17

```

	Query Match	25.0%; Score 169; DB 9; Length 524;
	Best Local Similarity	100.0%; Pred. No. 3.5e-73;
	Matches 169; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	237 ACCACCAACCACCATGGTCACACAGCGGSCCAGCGCTTGTGGTGCCGCAGAACCCAGGG	296

[illegible]

RESULT 7

US-10-125-540-76
; Sequence 76, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-540-76

RESULT 8

```

US-09-764-870-76
RESOL. 6
; Sequence 76, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764, 870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 2645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-76

```

Qy	616	AAA	675
Db	2582	AAA	2641
Qy	676	A A 676	
Db	2642	A 2642	

RESULT 9

```

US-09-832-129-21
; Sequence 21, Application US/09832129
; Publication No. US20030027297A1
; GENERAL INFORMATION:
; APPLICANT: Fiscellia et al.
; TITLE OF INVENTION: 19 Human secreted proteins
; FILE REFERENCE: P2045P1
; CURRENT APPLICATION NUMBER: US/09/832,129
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: PCT/US00/28664
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/163,085
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 60/172,411
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-129-21

```

DECEMBER 10

```

RESULT 10
US-09-948-820-38
; Sequence 38, Application US/09948820
; Publication No. US20030050460A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2034P1
; CURRENT APPLICATION NUMBER: US/09/948,820
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US/09/565,391
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US99/26409
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/108,207
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-948-820-38
Query Match          17.8%; Score 120; DB 9; Length 1457;
Best Local Similarity 100.0%; Pred. No. 3.2e-49;

```

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAA 616
Db 1331 TCAAA 1390
QY 617 AAA 676
Db 1391 AAA 1450

RESULT 11
US-10-239-676-106/c
; Sequence 106, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239, 676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 106
; LENGTH: 15832
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (2778, 2845..2846, 2865, 2916, 9325, 10589, 10612, 10809, 10817)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10821, 10827, 10830..10831, 10839, 12552, 13433, 15042)
US-10-239-676-106
Query Match 17.8%; Score 120; DB 9; Length 15832;
Best Local Similarity 100.0%; Pred. No. 2.5e-49;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 557 TCAAA 616
Db 4732 TCAAA 4673
QY 617 AAA 676
Db 4672 AAA 4613
RESULT 12
US-09-960-352-11058/c
; Sequence 11058, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalegan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11058
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 47-LIB34-055-Q1-EI-D12
US-09-960-352-11058

Query Match 17.6%; Score 119; DB 10; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 558 CAA 617
Db 154 CAA 95
QY 618 AAA 676
Db 94 AAA 36

RESULT 13
US-09-764-872-709
; Sequence 709, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 709
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-709

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 558 CAA 617
Db 17 CAA 76
QY 618 AAA 676
Db 77 AAA 135

RESULT 14
US-10-091-483-344/c
; Sequence 344, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-483-344

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 CAAA 617
|||||
Db 149 CAAA 90
|||||

QY 618 AA 676
|||||
Db 89 AA 31
|||||

RESULT 15
US-09-764-891-7491
; Sequence 7491, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7491
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7491

Query Match 17.6%; Score 119; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 CAAA 617
|||||
Db 17 CAAA 76
|||||

QY 618 AA 676
|||||
Db 77 AA 135
|||||

Search completed: June 16, 2003, 10:24:24
Job time : 155.802 secs

/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
BASE COUNT 172 a 193 c 183 g 100 t
ORIGIN

Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;
QY 67 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTGGACTACTGTCGGAGGA 126
Db |||||||
60 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTGGACTACTGTCGGAGGA 119
QY 127 GGTGAGTGTGCGGCGCGGCTTCCGCGGCACGCGCGCGGCTTCCAGCAGCAGCT 186
Db |||||||
120 GGTGAGTGTGCGGCGCGGCTTCCGCGGCACGCGCGCGGCTTCCAGCAGCAGCT 179
QY 187 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 243
Db |||||||
180 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 239
QY 244 CCACCATGCTTCCACACGCGCGGCTTCTGTTGCGCGGAGACGAGGTTCGAGGA 303
Db |||||||
240 CCACCATGCTTCCACACGCGCGGCTTCTGTTGCGCGGAGACGAGGTTCGAGGA 299
QY 304 GGACATCAACACCTGACCGCGGAGTTCACGAGCGGAGGAGTTCCTCGCGAGGGC 363
Db |||||||
300 GGACATCAACACCTGACCGCGGAGTTCACGAGCGGAGGAGTTCCTCGCGAGGGC 359
QY 364 TAACAGAGCGCGCGCGGCGGCTTCCAGCGCGGCTTCTGTTGCGCGGAGTTCGTTATG 423
Db |||||||
360 TAACAGAGCGCGCGCGGCGGCTTCCACGCGCGGCTTCTGTTGCGCGGAGTTCGTTATG 419
QY 424 TATGCTGTGTTGACTGTGTCAGGCTATCGTACGCTTCTGTTGCGCGGAGTTCGTTATG 483
Db |||||||
420 TATGCTGTGTTGACTGTGTCAGGCTATCGTACGCTTCTGTTGCGCGGAGTTCGTTATG 479
QY 484 CTCAGCTCTGTACGAATTACGACAAATAGCTCGTACGCTGATGATGATGATGATGATG 543
Db |||||||
544 ACTAATACCTACATCAAAAAA 610
604 AAAAAA 610
600 AAAAAA 606

RESULT 2
B0619315
LOCUS
DEFINITION
RN050606_SK.abl Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RN050606_SK.abl similar to No homology, mRNA
sequence.
B0619315
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Wang, H. and Bohnert, H. J.
Genetics of plant stress tolerance
Unpublished (2002)
Contact: Mark Fredrickson
Department of Plant Biology

University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
1. 648
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="RN050606_SK.abl"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

FEATURES
source

BASE COUNT 172 a 183 g 100 t
ORIGIN
Query Match 75.2%; Score 508.6; DB 14; Length 648;
Best Local Similarity 96.9%; Pred. No. 1.5e-46;
Matches 530; Conservative 0; Mismatches 14; Indels 3; Gaps 1;
QY 67 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTGGACTACTGTCGGAGGA 126
Db |||||||
60 GACACCAAGCGTGTGACCAATGGCTTACTACAGAGGTGGACTACTGTCGGAGGA 119
QY 127 GGTGAGTGTGCGGCGCGGCTTCCGCGGCACGCGCGCGGCTTCCAGCAGCAGCT 186
Db |||||||
120 GGTGAGTGTGCGGCGCGGCTTCCGCGGCACGCGCGCGGCTTCCAGCAGCAGCT 179
QY 187 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 243
Db |||||||
180 CGTCAAGGAGAAGTTCGAGGAGTTCGACACGCTATCAGCGCGCGCGCA---ACCACCA 239
QY 244 CCACCATGCTTCCACACGCGCGGCTTCTGTTGCGCGGAGACGAGGTTCGAGGA 303
Db |||||||
240 CCACCATGCTTCCACACGCGCGGCTTCTGTTGCGCGGAGACGAGGTTCGAGGA 299
QY 304 GGACATCAACACCTGACCGCGGAGTTCACGAGCGGAGGAGTTCCTCGCGAGGGC 363
Db |||||||
300 GGACATCAACACCTGACCGCGGAGTTCACGAGCGGAGGAGTTCCTCGCGAGGGC 359
QY 364 TAACAGAGCGCGCGCGGCGGCTTCCAGCGCGGCTTCTGTTGCGCGGAGTTCGTTATG 423
Db |||||||
360 TAACAGAGCGCGCGCGGCGGCTTCCACGCGCGGCTTCTGTTGCGCGGAGTTCGTTATG 419
QY 424 TATGCTGTGTTGACTGTGTCAGGCTATCGTACGCTTCTGTTGCGCGGAGTTCGTTATG 483
Db |||||||
420 TATGCTGTGTTGACTGTGTCAGGCTATCGTACGCTTCTGTTGCGCGGAGTTCGTTATG 479
QY 484 CTCAGCTCTGTACGAATTACGACAAATAGCTCGTACGCTGATGATGATGATGATGATG 543
Db |||||||
480 CTCAGCTCTGTACGAATTACGACAAATAGCTCGTACGCTGATGATGATGATGATGATG 539
QY 544 ACTAATACCTACATCAAAAAA 610
604 AAAAAA 610
600 AAAAAA 606

RESULT 3
B0619318
LOCUS
DEFINITION
RN050606D01_SK.abl Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RN050606D01_SK.abl similar to No homology, mRNA
sequence.
B0619318
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)